EXPLORING PEPTIDES IN THE BRAIN BY *IN VIVO* MICRODIALYSIS-CAPILLARY LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY

By

WILLIAM E. HASKINS

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

AUGUST 2003

To Lucas and Ksenija.

ACKNOWLEDGMENTS

My heart goes to those who gave me their love and encouragement during these difficult and wondrous years. It is my honor to thank my two greatest mentors, my advisor Dr. Robert T. Kennedy and Dr. David H. Powell; and my greatest colleague, Mr. Christopher J. Watson, for helping me do something I can be proud of, and for being patient with me. I also thank the members of the Kennedy Group, UF Spectroscopic Services, UF Chemistry; and my doctoral committee—Dr. Ben Horenstein, Dr. Satya Kalra, and Dr. Richard A. Yost—for their insight. I thank Alfred Chung (UF Protein Core) for peptide synthesis; Peter Viberg (Lund University) for help in emitter production; Weihong Tan (UF Chemistry) for use of a laser pipette-puller; and Maggie Stolls, Roger Reep, Tom Vickroy (UF Physiological Sciences) and Joanna Peris (UF Pharmacology) for confirming microdialysis probe placement. I also thank James Woods (U of Michigan Pharmacology) for conducting opioid receptor assays. I gratefully acknowledge the Analytical Chemistry Division at the University of Florida for a Laitinen Fellowship; the National Institutes of Health (NS38476); the National Science Foundation; and Bristol-Myers Squibb for funding this work. Lastly, I acknowledge my postdoctoral mentors, Dr. Ronald L. Hayes, Dr. Kevin K. Wang, and Dr. Nancy D. Denslow for their patience with me while I wrote this thesis and began exploring traumatic brain injury at the McKnight Brain Institute.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

LIST OF FIGURES vi

LIST OF TABLES ix

LIST OF ABBREVIATIONS x

ABSTRACT xii

CHAPTER

1 INTRODUCTION 1
  - Synpatic Release of Neuropeptides 1
  - *In Vivo* Microdialysis 3
  - Capillary Liquid Chromatography 5
  - Electrospray Ionization 7
  - Tandem Mass Spectrometry 9
  - Multi-Dimensional Techniques 10
  - Dissertation Overview 12

2 MONITORING and DISCOVERING ENDOGENOUS NEUROPEPTIDES BY *IN VIVO* MICRODIALYSIS-CLC-MS^2 13
  - Introduction 13
  - Experimental 15
  - Results and Discussion 19
  - Conclusion 40

3 DISCOVERING ENDOGENOUS NEUROPEPTIDES BY *IN VIVO* MICRODIALYSIS-CLC-DATA-DEPENDENT-MS^2 41
  - Introduction 41
  - Experimental 43
  - Results and Discussion 50
  - Conclusion 73
  - Supporting Information 75

4 STUDYING ENDOGENOUS NEUROPEPTIDE PROCESSING BY *IN VIVO* MICRODIALYSIS-CLC-MS^2 82
Introduction 82
Experimental 84
Results and Discussion 84
Conclusion 105

5 SUMMARY AND FUTURE DIRECTIONS 106
Summary 106
Future Directions 111

APPENDIX 115
A EXAMPLE OF CLC-MS^2 CALCULATIONS 115
B EXAMPLE OF PEPTIDE SEQUENCING AND PROTEIN PRECURSOR IDENTIFICATION. 118
C TROUBLESHOOTING THE *IN VIVO* MICRODIALYSIS-CLC-MS^2 SYSTEM 124

LIST OF REFERENCES 125

BIOGRAPHICAL SKETCH 140

LIST OF FIGURES

Figure Page

1-1. Synaptic release of neuropeptides............1

1-2. A model of proteolytic processing for PEA............3

1-3. *In vivo* microdialysis scheme............4

1-4. Photograph of *in vivo* microdialysis of a freely moving rat............5

1-5. Scanning electron micrograph of a 25 μm -i.d CLC column packed with 5 μm octadecylsilane (C18) particles for preconcentration, desalting and separation of attomole amounts of peptides in microliter sample volumes by reverse-phase gradient-elution 6

1-6. Photograph of electrospray ionization............8

1-7. Peptide sequencing by collision-induced dissociation............9

1-8. Flow chart for our multi-dimensional method, *in vivo* microdialysis-CLC-MS^2, developed for monitoring and discovering peptides in the brain............11

2-1. Block diagram of the automated two-pressure CLC-MS^2 system with on-line microdialysis sampling............20

2-2. CLC columns with integrated electrospray emitters............22

2-3. Effect of gradient steepness parameter (b) on resolution (R) and sensitivity (peak height) as a function of flow rate............24

2-4. Effect of gradient steepness parameter (b) on resolution (R) for different column lengths and flow rates (■ = 2 cm long column at 20 nL/min; ● = 10 cm long column at 50 nL/min; ♦ = 2 cm long column at 70 nL/min)............25

2-5. High sensitivity CLC-MS^2 measurement of a mixture of Met-enkephalin and Leu-enkephalin at 33 pM (1.8 μL injected corresponding to 59 amol of each peptide loaded onto the column) each in aCSF using a 2 cm column............27

2-6. Illustration of peptide carry-over in the injection valve and carry-over prevention, using deuterated standards 29

2-7. *In vivo* detection of Met- and Leu-enkephalin during basal and K^+-stimulated conditions 31

2-8. *In vivo* monitoring results for Met-enkephalin (■), Leu-enkephalin(●) and an unknown peptide (▲) for a single rat 34

2-9. *In vivo* identification of peptides by CLC-data-dependent-MS^2..36

2-10. Sequencing of a novel peptide by time-segmented MS^2 from sample collected *in vivo* during K^+-induced depolarization 37

2-11. Confirmation of the sequence of a novel peptide by time-segmented MS^3 from sample collected *in vivo* during K^+-induced depolarization 38

2-12. Amino acid sequence of PEA..39

3-1. Triplcate, 0.4 µL, injections of a tryptic digest of lysozyme, carbonic anyhdrase, and conalbumin at 10 nM with 1 nM $YGGFD_5L$ (400 amol injected on-column)............55

3-2. *In vivo* microdialysis-CLC-MS^2 analysis of CSF collected from the rat striatum during K^+-induced release conditions 56

3-3. Comparison of total ion chromatograms (TIC) and reconstructed ion chromatograms (RICs) for data-dependent MS^2 of dialysate collected from rat striatum during basal (B and D, respectively) and depolarizing (A and C, respectively) conditions showing the most abundant precursor ions observed *in vivo*..57

3-4. Representative MS^2 spectra and sequences for peptides produced from proteolytic processing of PEA (A), neurogranin (B), fibrinogen α (C), fibrinogen β (D), excitatory amino acid transporter (E), and brain acidic membrane protein (F)..63

3-5. MS^2 spectra (A and C) from the *in vivo* library and the difference MS^2 spectrum (B) for SPQLEDEAKE observed in two rats..66

3-6. A summary of the *in vivo* results illustrating the six step data-reduction strategy described in the experimental section 67

3-7. Diagram showing the interaction of multiple selection criteria and the overlap between MS^2 spectra remaining at data-reduction steps 4 and 5 (see experimental section). 72

3-8. MS^2 spectra and sequences for peptides produced from proteolytic processing of PEA (A), neurogranin (B), fibrinogen α (C), fibrinogen β (D), excitatory amino acid transporter (E) and brain acidic membrane protein (F) 75

4-1. *In vivo* microdialysis-CLC-MS^2 analysis of CSF collected from the rat striatum during depolarization-induced release of peptides into the ECF 85

4-2. Amino acid sequence of PEA and PEA-derived peptides discovered in this work93

4-3. Non-exhaustive depiction of selected PEA-derived peptides by amino acid position. 99

4-4. Hypothetical intermediates (HIs) of PEA processing and novel PEA-processing observed *in vivo* 100

4-5. *In vivo* effect of peptide I_{1-10} on neurotransmitter levels 101

5-1. *In vivo* microdialysis-CLC-MS^2 base peak reconstructed ion chromatograms (RICs) show the most abundant product ions observed during sleeping (A) and prolonged-wakefulness (B) states. 113

B-1. Sequest summary report for a tryptic digest of cytochrome c 118

B-2. Sequest cyctochrome c report for a tryptic digest of cytochrome c 119

B-3. Mascot summary report for a tryptic digest of cytochrome c 120

B-4. Mascot cytochrome c report for a tryptic digest of cytochrome c 121

B-5. Mascot annotated MS^2 spectrum for the tryptic peptide TGPNLHGLFGR from the tryptic digest of cytochrome c 122

B-6. Lutefisk report for the tryptic peptide TGPNLHGLFGR from the tryptic digest of cytochrome c 123

LIST OF TABLES

Table Page

3-1. Summary of *in vivo* microdialysis-CLC-MS^2 results 58

3-2. Peptides and proteolytic processing patterns observed *in vivo*..........59

3-3. Summary of *in vitro* CLC-MS^2 results 80

3-4. Peptides and proteolytic processing patterns observed *in vitro*..........81

4-1. Putative proteases involved in the proteolytic processing patterns observed *in vivo* 88

A-1. Example of CLC-MS^2 calculations 115

C-1. Symptoms, diagnosis and treatment for common problems in each dimension of the *in vivo* microdialysis-CLC-MS^2 system. 124

LIST OF ABBREVIATIONS

aa	Amino Acid
aCSF	Artificial cerebral spinal fluid
amol	Attomole
amu	Atomic mass units
BAM	Bovine adrenal medulla
BAMP	Brain acidic membrane protein
CaM	Calmodulin
CLC	Capillary liquid chromatography
EAAT1	Excitatory amino acid transporter 1
ECF	Extracellular fluid
HI	Hypothetical intermediate
HPLC	High-performance liquid chromatography
id	Internal diameter
LC	Liquid chromatography
μL	Microliter
MS	Mass spectrometry
MS^2	Tandem mass spectrometry
MWCO	Molecular weight cut-off
NG	Neurogranin

od	Outer diameter
PAR	Protease activated receptor
PEA	Preproenkephalin A
PC1/3	Proprotein convertase 1 and 3
PC2	Proprotein convertase 2
PKC	Protein kinase C
pM	Picomolar
QIT	Quadrupole ion trap
RIA	Radioimmunoassay
RIC	Reconstructed ion chromatogram
RSD	Relative standard deviation
SD	Standard deviation
SEM	Standard error of the mean
S/N	Signal-to-noise ratio
SV	Secretory vesicle
TIC	Total ion chromatogram
UV	Ultraviolet

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

EXPLORING PEPTIDES IN THE BRAIN BY *IN VIVO* MICRODIALYSIS-CAPILLARY LIQUID CHROMATOGRAPHY-TANDEM MASS SPECTROMETRY

By

William E. Haskins

August 2003

Chair: Robert T. Kennedy
Major Department: Chemistry

Neuropeptides are hormones, neurotransmitters, and neuromodulators important in learning, memory, appetite regulation, sensory perception and disease. Many recent studies have shown that post-translational proteolytic processing produces neuropeptides that were not predicted from known enzymatic cleavage sites. Thus, the mechanism for regulating the bioactivity and stoichiometry of peptide products from protein precursors remains a mystery. In this work is described the development and application of novel bioanalytical methods and tools for monitoring and discovering endogenous neuropeptides in the mammalian brain.

Time-segmented and data-dependent tandem mass spectra were collected from male Sprague-Dawley rats during basal and K^+-induced depolarizing conditions. A 4 mm microdialysis probe sampled the globus pallidus or striatum region of the rat brain at 0.6 μL/min for on-line, *in vivo* microdialysis-capillary liquid chromatography-tandem mass spectrometry. Multiple peptides at low attomole levels (low picomolar concentrations) were quantified and sequenced in microdialysate samples by using 25 μm i.d. columns with integrated 3 μm i.d. electrospray emitters, optimized gradients and a quadrupole ion trap mass spectrometer.

This is the first report of on-line *in vivo* microdialysis of endogenous neuropeptides in combination with 1) separation and simultaneous monitoring, 2) data-dependent measurement, 3) mass spectral library construction and 4) novel peptide discovery. The discovery of endogenous neuropeptides provides new insights into inter-neuron chemical signaling. This method should be useful for monitoring known peptides, studying processing of endogenous peptides from protein precursors and characterizing novel peptides that are potentially bioactive. Moreover, exploring the extracellular release of endogenous neuropeptides suggests novel biomarkers for biomedical applications.

CHAPTER 1
INTRODUCTION

Synaptic Release of Neuropeptides

Peptide signaling in the central nervous system has been implicated in many physiological functions, including learning, memory, appetite regulation, and sensory perception[1]. Figure 1-1 is a cartoon of the synaptic junction during release of neuropeptides.

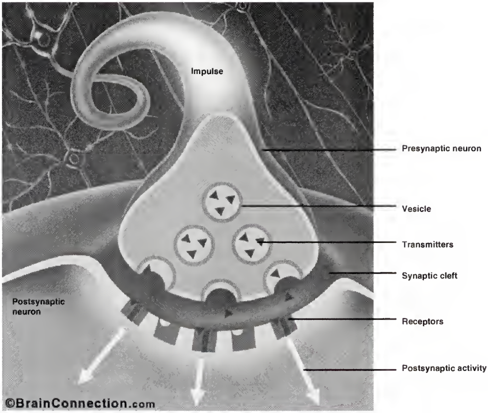


Figure 1-1. Cartoon of the synaptic junction during release of neuropeptides (adapted from http://www.brainconnection.com).

Following a neuronal impulse in the presynaptic neuron, vesicles containing neurotransmitters and neuromodulators fuse with the plasma membrane where they are released into the synaptic cleft, bind to receptors in postsynaptic neurons and diffuse into the extracellular fluid (ECF) to exert their actions farther from the synapse.

Neuropeptides are produced by the action of numerous intra- and extracellular proteases, endo- and exopeptidases, on precursor proteins[2]. Differential processing by various highly-specific proteases produces a complex mixture of peptide intermediates and bioactive peptides. Proteases typically cleave at mono- and di-basic amino acid residues (e.g., R, K, KK and RK) flanking the N- and C-terminus of the peptide sequence within the protein precursor. A model of proteolytic processing is shown in Figure 1-2 for preproenkephlin A (PEA)[3]. Examples of previously-observed, PEA-derived peptide intermediates (e.g., peptide I) and bioactive peptides (e.g., the neuropeptides Met- and Leu-enkephalin[4]) are shown by unshaded and shaded regions of the protein sequence, respectively.

Previous studies on proteolytic processing made no discrimination between intra- and extracellular-peptides and were conducted primarily in bovine adrenal medulla (BAM) cells. In this work, we developed a new multi-dimensional technique, *in vivo* microdialysis-capillary liquid chromatography (CLC)-tandem mass spectrometry (MS^2), and applied it to monitoring and discovering endogenous peptides in the extracellular space of the brain[5,6]. Key aspects of each dimension of the technique are described in the following paragraphs.

In Vivo Microdialysis

In vivo microdialysis enables dynamic measurements of the ECF in discrete regions of the brain with minimal damage to the surrounding tissue[7]. In this approach, the microdialysis probe mimics the action of a capillary blood vessel by passively sampling the ECF for molecules involved in cell signaling. As shown in Figure 1-3, the microdialysis probe samples analytes (A, e.g., neuropeptides for CLC-MS^2 analysis) from the ECF of the brain as they diffuse across the semi-permeable dialysis membrane.

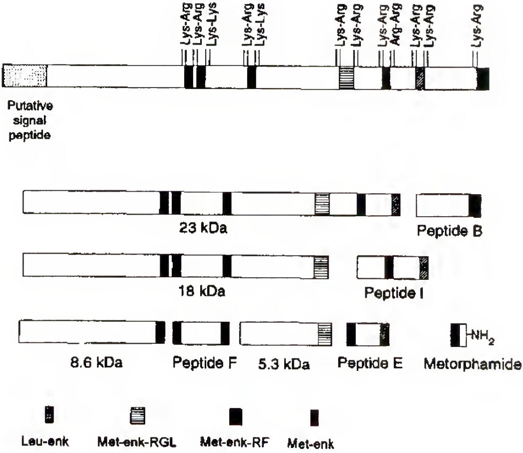


Figure 1-2. A model of proteolytic processing for PEA. Examples of intermediates and bioactive peptides are shown by unshaded and shaded regions of the protein sequence, respectively (adapted from reference 3).

Artificial cerebral spinal fluid (aCSF) is pumped through the probe inlet and outlet to collect fractions for off-line or on-line analysis. Molecules larger than the molecular weight cut-off (MWCO) of the membrane (P, e.g., proteins) remain in the ECF. Recovery of peptides through the microdialysis probe depends on a number of factors including the flow rate, MWCO, hydrophobicity of the peptide, adsorption and other interactions with the membrane, dead volume, dimensions of the probe and brain region sampled[7].

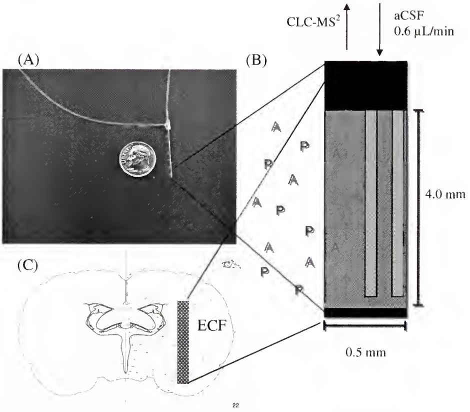


Figure 1-3. *In vivo* microdialysis scheme. (A) A photograph of the microdialysis probes used throughout this work with a dime shown for size comparison (Courtesy of David H. Powell). (B) A cartoon of the active region of the microdialysis probe. The probe samples analytes (A, e.g., neuropeptides for CLC-MS^2 analysis) from the ECF of the brain as they diffuse across the semi-permeable dialysis membrane. Molecules larger than the molecular weight cut-off (MWCO) of the membrane (P, e.g., proteins) remain in the ECF (Courtesy of Christopher J. Watson). (C) A cartoon of the active region of the microdialysis probe placed in the globus pallidus region of the brain (sagittal view of brain slice).

Behavioral, physiological and pharmacological experiments are enabled by coupling *in vivo* microdialysis to various analysis techniques. Figure 1-4 is a photograph of a freely-moving rat with a microdialysis probe implanted into the striatum region of its brain. Because the animal behaves normally despite the probe, behavioral studies can also be performed. For example, samples can be collected during sleeping and prolonged-wakefulness states to investigate the molecular correlates of sleep[8] (Chapter 5). Alternatively, *in vivo* microdialysis can be used to administer exogenous compounds to determine pharmacological effects. In this example, a drug that induces sleep can be infused into the probe while endogenous compounds or metabolites of the drug are sampled for analysis.

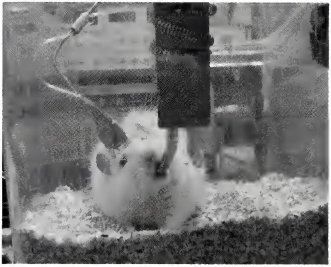


Figure 1-4. *In vivo* microdialysis of a freely-moving rat. (Courtesy of Christopher J. Watson).

Capillary Liquid Chromatography

Neuropeptides are present at attomole levels (1-100 pM) in microliter sample volumes (1-100 μL) of dialysate containing a complex mixture of salts and neurotransmitters. Thus, the analysis technique must be extremely selective, sensitive and rapid in order to maximize the flux of chemical information from the brain. In this work, we coupled *in vivo* microdialysis sampling with capillary liquid chromatography (CLC)-tandem mass spectrometry (MS^2) to both monitor and discover peptides found in the extracellular space of the brain. CLC offers several attractive features for separation of peptides in mass- and concentration-limited samples.

The concentration at the apex of a chromatographic peak (C_{max}) is inversely proportional to the square of the column internal diameter (d_c or i.d.) as given by the following approximation:

$$C_{max} \sim N^{1/2} / Ld_c^{2}(1+k) \tag{1-1}$$

where N is the number of theoretical plates, L is the column length and k is the capacity factor. Therefore, a 34,000 fold improvement in sensitivity is expected when going from

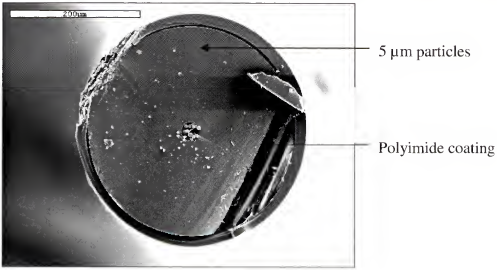


Figure 1-5. Scanning electron micrograph of a 25 µm -i.d, 375 µm -o.d. CLC column packed with 5 µm octadecylsilane (C18) particles for preconcentration, desalting and separation of attomole amounts of peptides in microliter sample volumes by reverse-phase gradient-elution. The outer ring of material is a protective polyimide coating that surrounds the fused silica capillary.

a conventional bore, 4.6 mm -i.d. column to a 25 μm -i.d[9] CLC column! Figure 1-5 shows a scanning electron micrograph of the 25 μm -i.d. CLC columns used throughout this work (prepared in-house).

The enormous gain in sensitivity combined with preconcentration, desalting and separation by reverse-phase gradient-elution enables attomole measurements of endogenous peptides from microliter sample volumes. Analytes may be preconcentrated and desalted in a weak mobile phase, then separated in an increasingly strong mobile phase. Separation of peptides by CLC depends on a number of factors, including the flow rate, hydrophobicity of the peptide, adsorption and other interactions with the CLC system, dead volume, dimensions of the CLC column, mobile phase, stationary phase and particle size[10].

Capillary liquid chromatography has been coupled to a number of selective detectors for analysis of mass- and concentration-limited samples[10]. This includes laser-induced fluorescence (LIF), electrochemical detection (EC) and tandem mass spectrometry (MS^2). The most selective of these, MS^2, was traditionally difficult to couple to CLC. However, recent improvements in electrospray ionization (ESI)[11] have enabled CLC-MS^2 to become a mainstay of proteomics methods.

Electrospray Ionization

Electrospray ionization[12], provides a sensitive interface between CLC and MS^2 [13]. A bright field image illustrating the ESI process is shown in Figure 1-6. In this work, a potential difference is applied between the CLC column and the heated capillary of the quadrupole ion trap (QIT) mass spectrometer to generate a Taylor cone of charged droplets undergoing fission and solvent evaporation. Multiply-charged peptide ions (e.g., $[M+nH]^{n+}$ for positive ion mode) are formed in this process and transferred to the mass spectrometer for analysis. Ionization and ion transfer efficiency depend on a number of factors, including the flow rate, mobile phase, hydrophobicity and gas phase basicity of the peptide, shape of the needle tip, distance between the CLC column and the QIT and the potential difference applied[11,12].

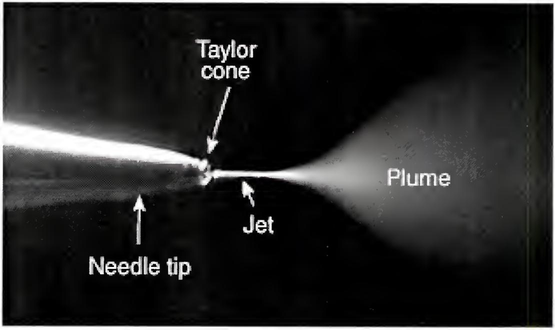


Figure 1-6. Photograph of electrospray ionization.

Peptides and salts compete for charge during ESI; therefore, desalting by CLC proved to be very useful for ESI of peptides in salty samples such as dialysate. Our method for dialysis analysis is described as *in vivo* microdialysis-CLC-MS^2 instead of *in vivo* microdialysis-CLC-ESI-MS^2, although the latter is more accurate, because the other popular ionization process, matrix-assisted laser/desorption ionization (MALDI), has not yet been efficiently coupled to CLC[14]. Once the peptides in dialysate have been preconcentrated, desalted and separated by CLC, ionized and transferred to the mass spectrometer by ESI, they can be quantified and sequenced by MS^2.

Tandem Mass Spectrometry

Tandem mass spectrometry (MS^2) is ideally suited for monitoring and discovering endogenous peptides because charge migration and collision-induced dissociation (CID) proceed along the peptide backbone in a manner that is sequence-specific[15]. In peptide sequencing by MS^2, gas-phase precursor ions are isolated, fragmented by CID and the mass-to-charge ratio (*m/z*) of the product ions are measured (see Figure 1-7). The sequence of a peptide can be interpreted from the pattern of product ions in its MS^2 spectrum. However, this is a difficult and time-consuming process; therefore, database-searching programs are now used to compare MS^2 spectra predicted from protein databases with MS^2 spectra observed experimentally[16]. In this way, peptide sequences and protein precursors can be identified automatically. Alternatively, two or more product ions may be monitored simultaneously for sequence-specific quantification.

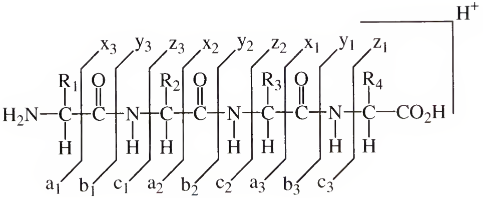


Figure 1-7. Peptide sequencing by collision-induced dissociation. Product ions, arising from cleavage of the precursor ion's peptide backbone and charge migration, retain charge on either the N-terminus (a-, b- or c-type ions) or the C-terminus (x-, y- or z-type ions).

Multi-dimensional Techniques

Multi-dimensional, or hyphenated, techniques (e.g., LC-GC, LC-MS, LC-MS-NMR) provide maximum resolution and informing power for complex mixtures provided that each orthogonal dimension improves the selectivity and sensitivity of the technique[17]. However, one limitation of these approaches is that the transfer of analyte from one dimension to the next must be efficient in order to achieve the mininum signal necessary for analysis. Mass- and concentration-limited samples such as dialysate further exacerbate this problem.

A flow chart for the multi-dimensional method, *in vivo*-microdialysis-CLC-MS^2, developed for monitoring and discovering peptides in the brain is shown in Figure 1-8. In discovery applications, proteomics-based methods were employed to correlate endogenous peptides with protein precursors across multiple animals. CLC-MS^2 chromatograms containing thousands of MS^2 spectra were subjected to data-reduction, database searching and sequence validation. Significant MS^2 spectra were assigned to peptide sequences and protein precursors by comparing basal and stimulated conditions. In monitoring applications, endogenous peptides were quantified with sequence-specificity and correlated to specific time-points during basal and stimulated conditions.

Losses of analyte between each dimension of the method were expected; however, these were minimized by miniaturization approaches, including reducing the dead volume between the microdialysis probe and the CLC column, reducing the i.d. of the CLC column, reducing the i.d. of the electrospray emitter, reducing the dead volume between the CLC column and the electrospray emitter and reducing the distance between the electrospray emitter and the mass spectrometer. The following chapters describe success at monitoring and discovering peptides in the brain.

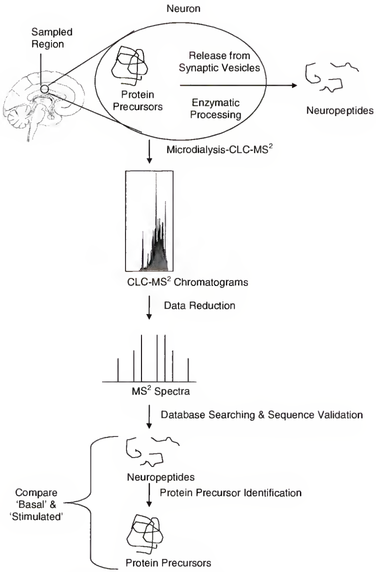


Figure 1-8. Flow chart for our multi-dimensional method, *in vivo* microdialysis-CLC-MS^2, developed for monitoring and discovering peptides in the brain. Neuropeptide synthesis in the brain by enzymatic processing of protein precursors released from synaptic vesicles is followed by collection of MS^2 spectra at basal and stimulated levels, peptide sequencing and protein precursor identification.

Dissertation Overview

The overall objective of this research is to develop, evaluate and apply a method for monitoring and discovering peptides in the brain of living animals.

Chapter 2 describes the development, evaluation and application of a novel *in vivo* microdialysis-capillary liquid chromatography (CLC)-tandem mass spectrometry (MS^2) system for monitoring peptides. Simultaneous, on-line monitoring of Met-enkephalin, Leu-enkephalin and a previously unknown peptide at attomole levels with 30-min temporal resolution is described. Preliminary evaluation of the system for discovering previously unknown peptides by data-dependent MS^2 is also investigated. Chapter 3 describes the development, evaluation and application of a proteomics-based data-reduction strategy in conjunction with the *in vivo* microdialysis-CLC-data-dependent-MS^2 system for discovering previously unknown peptides. The reproducibility of peptide sequencing and protein precursor identification at attomole levels is also investigated. Chapter 4 describes the biological significance of the peptide sequences and protein precursors discovered in this work. Proteolytic processing is described as a function of the putative proteases involved in the cleavage of each peptide sequence from its protein precursor. In chapter 5, a summary of this work and future directions are presented.

CHAPTER 2
MONITORING ENDOGENOUS NEUROPEPTIDES BY *IN VIVO* MICRODIALYSIS-CLC-MS[2]

Introduction

Neuropeptides are an important and diverse class of interneuron-signaling molecules that act as hormones, neuromodulators, and neurotransmitters. Peptide signaling in the central nervous system has been implicated in many physiological functions, including learning, memory, appetite regulation, and sensory perception. Although direct measurement of extracellular peptide levels *in vivo* is invaluable in elucidating peptide function, more indirect methods such as the effect of pharmacological agents or the post-mortem analysis of tissue for peptides or mRNA are typically used in studies of neuropeptides. These indirect methods are popular because sampling techniques for *in vivo* monitoring, such as microdialysis, produce microliter samples with amol quantities of peptide, thus making analytical measurements difficult.

When *in vivo* peptide monitoring is performed, the analytical technique of choice is radioimmunoassay (RIA) because it has excellent specificity and, with appropriate optimization, can achieve detection limits of 100 amol in microliter fractions[18]. Despite the high specificity, RIA results can be ambiguous due to cross-reactivity of the labeled antibody with similar peptides[19,20]. HPLC combined with RIA of fractions can minimize specificity limitations; however, this combination is a cumbersome procedure that is usually avoided. RIA's are also usually limited to measurement of only a single peptide per fraction. Thus, relationships between different peptides and stimuli must be inferred from data collected from different animals. Finally, RIA does not allow monitoring or discovery of novel neuropeptides. In the past, discovery of novel neuropeptides was accomplished by preparative scale isolation of bioactive components, purification, and Edman sequencing.

Newer approaches to detection of endogenous neuropeptides in dialysate, including CLC with electrochemical detection (CLC-EC)[21,22] and tandem mass spectrometry (MS^2)[11,23,24] have been reported recently. MS^2 is attractive because it offers the possibility of detecting peptides with sequence specificity and can be used, in principle, for any peptide. In the seminal work on MS^2 for *in vivo* peptide detection, 10 µL microdialysis samples were preconcentrated and desalted on 50 µm i.d. columns coupled to a triple quadrupole mass spectrometer by a microelectrospray interface[11,24]. Endogenous Met-enkephalin[23] and neurotensin[24] were measured in microdialysate fractions using the selected-reaction-monitoring (SRM) scan mode, where only one or a few precursor-to-product ion transitions are monitored over a narrow scan range (e.g., 2 *m/z*). This method was revolutionary in its use of microelectrospray for high sensitivity; however, the use of step-gradients and SRM precluded simultaneous multi-analyte monitoring or characterization of unknown peptides.

In order to improve upon this approach for *in vivo* microdialysis experiments, we have explored the use of: 1) miniaturized LC columns and electrospray emitters for better sensitivity, 2) optimized mobile phase gradients to allow for peptide separations and multi-analyte monitoring, and 3) a quadrupole ion trap (QIT), which provides full-scan MS^2 at a sensitivity comparable to SRM experiments on triple quadrupole instruments[25].

In addition to improving confidence in the measurement of known peptides[26], full-scan MS^2 offers the possibility of quantifying and characterizing novel peptides by sequencing at amol levels[27].

Experimental

Chemicals and Reagents. CLC solvents were purchased from Burdick and Jackson (Muskegon, MI). Peptides were from Sigma (St. Louis, MO). Acetic acid and hydrofluoric acid were purchased from Fisher Scientific (Pittsburgh, PA). Ethanol was purchased from J.T. Baker (Phillispburg, NJ). The following chemicals were purchased from Aldrich (Milwaukee, WI) to prepare macroporous photopolymer frits: isooctane, toluene, trimethylolpropane trimethacrylate, glycidyl methacrylate, and benzoin methyl ether. Artificial cerebral spinal fluid (aCSF) used for microdialysis perfusion consisted of 145 mM NaCl, 2.68 mM KCl, 1.10 mM $MgSO_4$, and 1.22 mM $CaCl_2$. The high K^+, low Na^+ perfusate solution for depolarization experiments consisted of 145 mM KCl and 2.62 mM NaCl with other salts the same as aCSF. Mobile phases and aCSF were prepared weekly, and were filtered with 20 nm-pore size aluminum oxide filters (Fisher) to remove particulates.

***In Vivo* Microdialysis.** Commercial microdialysis probes (CMA/10 with 20,000 Da molecular weight cut-off, CMA, Solna, Sweden) with 4-mm active lengths were used as described previously by RIA[18], MS-MS[23,24,24], and CLC-EC[21,22,22]. The outlet tubing of the microdialysis probe was replaced with fused silica capillary (50 μm i.d., 360 μm o.d., Polymicro Technologies, Phoenix, AZ) to minimize dead volumes. Prior to use, microdialysis probes were rinsed with 70% ethanol at a flow rate of 0.6 μL/min per the manufacturer's instructions. Deuterated Leu-enkephalin (UF Protein Core), where 5 non-exchangeable deuterium atoms were incorporated into the phenylalanine ring, was used to calculate the relative recovery through the microdialysis probe. *In vitro* relative recovery of 600 pM deuterated Leu-enkephalin in aCSF was 51 ± 16% (n = 4) at 37 °C. Surgical preparation and procedures were the same as previously described[21]. Correct placement of the microdialysis probe was confirmed by inspection of cresyl violet-stained tissue sections cut by a cryomicrotome.

Preparation of CLC Columns with Integrated Electrospray Emitters. To prepare columns, a 1 cm section of polyimide was removed from a 20 cm length of 25 μm i.d. fused silica with a flame. The capillary was loaded with a solution of benzoin methyl ether (1.5% w/w), glycidyl methacrylate (7% v/v) and trimethylolpropane trimethacrylate (16% v/v) in a 3:1 isooctane:toluene solvent, using He pressure, and was placed inside PEEK sleeves (Upchurch Scientific, Oak Harbor, WA) that formed a mask that exposed a ~100 μm length of the bare fused silica. Polymerization was achieved by illuminating the exposed region of the capillary with a UV lamp (Spectronics, Westbury, NY) for 30 min[28,29]. (In practice, 10 capillary frits were prepared simultaneously). Frits were flushed with acetone, and were dried for storage by passing He through the capillaries for 5 min at 500 psi. A P-2000 CO_2 laser puller (Sutter Instruments, Novato, CA) was used to create integrated electrospray emitters less than 1 cm beyond the frit. The settings used on the P-2000 were heat 350, filament 0, velocity 20, delay 135, and pull 125. The program was cycled 3-4 times to generate pulled capillary columns with less than 1 μm orifices. The resulting pulled columns were etched with 50% hydrofluoric acid (use extreme caution and neutralize with $CaCl_2$)[30] for 30 s to create sharp-edged electrospray emitters with approximately 3 μm i.d. Pulled columns were packed with an acetone slurry (10 mg/mL) of 5 μm Alltima C18 reversed-phase particles (Alltech, Deerfield, IL) at 1000 psi as described elsewhere[9]. Only 2 cm of the total 13 cm length of fused silica was packed, unless stated otherwise. The void at the head of the column did not contribute to extra-column band-broadening because all samples were injected in weak mobile phases to allow the analytes to stack at the head of the column.

Automated Two-Pressure LC-MS^2 System and Operation. To achieve rapid switching between high-flow rates desired for sample loading and low-flow rates desired for chromatographic separation, a two-pressure system was used, as shown in Figure 2-1. The system utilized two six-port valves (C2 valves, Valco, Houston, TX): valve 1 was used to select the pump for separation or sample loading and valve 2 was the injection valve. To perform an injection, samples were loaded onto the sample loop of the injector valve (valve 2) either by a conventional syringe port or from the microdialysis probe, as shown in Figure 2-1. After filling the sample loop, the splitter at the waste port (W) was closed and valve 2 actuated to allow sample to be loaded onto the column. The loading pump (Model DSFH-151, Haskel, Burbank, CA) was selected by valve 1 (flow path shown in Figure 2-1) and sample pumped onto the column at approximately 370 nL/min. After the sample was loaded, valve 2 was actuated to remove the sample loop from the flow path and the column rinsed for 5 min at 370 nL/min to desalt the column. Valve 1 was then used to select the gradient syringe pump (Model 100DM, ISCO, Lincoln, NE), and gradient elution was initiated with a post-valve split from 4 μL/min to 20 nL/min unless stated otherwise. Splitter 1 was open and splitter 2 (Figure 2-1) was closed during separation to minimize the dwell time of the gradient. Mobile phase A contained 1% acetic acid in water and mobile phase B contained 1% acetic acid in methanol. The gradients were changed as described in the text. The solvent reservoir of the sample-loading pump contained 1% acetic acid in water.

A personal computer (model E-4200, Gateway, North Sioux City, SD) and I/O board (AT-M10-16XE-50, National Instruments, Austin, TX) was used to control the timing of the system via a Labview (National Instruments) program written in-house. The program independently controlled the injection time, column rinse time, separation/electrospray time, and re-equilibration time. All measurements were made with the following CLC parameters, unless specified otherwise: preconcentration time = 5 min (1.8 μL), desalting time = 5 min (1.8 μL), separation/electrospray time = 15 min, re-equilibration time ≤ 5 min. In addition to valve control, the program triggered a pulse of N_2 gas to remove salt droplets from the electrospray emitter before the emitter was positioned at the entrance of the mass spectrometer. The pulse of N_2 prevented salt contamination of the ion optics by aCSF solutions that accumulated on the tip of the column during sample loading. The tip of the column was automatically positioned at the inlet of the mass spectrometer by a servomotor-driven translation stage and motion controller/driver after injection and rinsing (CMA 12CCCL and ESP 300, Newport). Once the emitter was in position, the gradient syringe pumps and the QIT (LCQ-Deca, ThermoFinnigan, San Jose, CA) were triggered to begin separation and electrospray. Electrospray voltage (1.5 kV) was applied at a liquid junction downstream of the gradient splitter tee to prevent contamination by oxidation products. PEEK spacers were constructed in-house to electrically insulate the valves from the actuators (Valco).

All measurements were made with the following QIT parameters, unless specified otherwise: automatic gain control (AGC) on, max AGC time = 300 msec, q = 0.25, isolation width = 3 amu, normalized collision energy = 35%, activation time = 0.25 msec, and the default number of microscans and target count values. Sequest software was used to correlate uninterpreted MS^2 data with protein databases for automated peptide sequencing and protein precursor identification. Automatic tuning of the QIT ion optics was accomplished by loop injection of an 18 nM solution of neurotensin 1-11 (qLYENKPRRPYIL, where q = pyroglutamate) in 50% methanol:1% acetic acid after manual positioning of the emitter less than 0.5 mm from the entrance to the heated capillary (150 C). A video microscope (VZM 1000, Edmund Scientific, Barrington, NJ) was used to monitor the position of the emitter and electrospray stability at 800X magnification. Once tuning was completed, excess neurotensin 1-11 was eluted from the column by ramping the gradient from 50 to 90% B. Splitter 1 was closed and splitter 2 (S1 and S2, respectively, in Figure 2-1) was open during tuning to prevent splitting the sample while maintaining a flow rate equivalent to the separation flow rate.

Results and Discussion

CLC Columns with Integrated Electrospray Emitters. Previous work for high sensitivity detection of neuropeptides in dialysates utilized 50 μm i.d. columns with 50 μm i.d. integrated electrospray emitters operated at flow rates of 200-800 nL/min[11,23,24]. Reducing the column/emitter dimensions and the separation/electrospray flow rate were expected to improve separation efficiency, mass sensitivity, and concentration sensitivity.

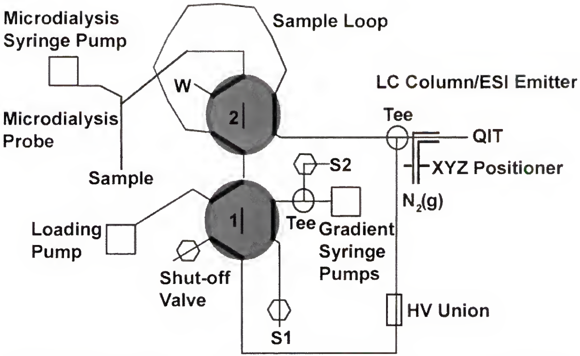


Figure 2-1. Block diagram of the automated two-pressure CLC-MS^2 system with on-line microdialysis sampling. Valve 1 is used for pump selection and valve 2 is the injection valve. S1 and S2 are splitter capillaries (1 m x 50 µm i.d.) that can be shut off with valves (hexagons) as shown. W is a waste port. HV is the stainless steel union where high voltage is applied to generate electrospray. The flow path shown is for sample loading onto the column. Operation of the system is described in the text.

Twenty-five µm i.d. columns packed with 5 µm particles and 3 µm i.d. electrospray emitters were used to gain substantial miniaturization but retained routine preparation and operation. Figure 2-2 illustrates a column with integrated electrospray emitter (2A), the macroporous frit (2B), and the orifice at the electrospray emitter tip (2C). The emitters could be reproducibly made with a 3.1 ± 0.2 µm i.d. (n = 5). Frits were prepared by *in situ* photo-polymerization[28,29], instead of sintering particles[9], because this procedure allowed the frits to be readily placed near the end of the column to minimize extra-column band-broadening. (Only a 2% loss in resolution was observed for frits placed ~ 1 cm from the electrospray emitter orifice in 25 µm i.d. columns operated at separation flow rates 20 nL/min.) In addition, such frits were found to be more reliable than sintered particle frits in that they prevented unsintered particles from clogging the electrospray emitter. The main limitation of small columns and tips was that, after approximately 25 injections of high volume (1.8 μL) aCSF solutions, clogging of the electrospray emitter by salt deposits caused a decrease in flow rate. Preparing columns daily circumvented clogging problems.

In addition to the column/tip dimensions, a second important feature of this system was use of an automated two-pressure LC system, which greatly reduced analysis time. This system enabled use of high-flow rates to minimize the sample loading time and column rinse time. Thus, 1.8 μL samples could be loaded in 5 min at 370 nL/min (3200 psi). By switching to a second pump that was already equilibrated to a lower pressure after loading, the separation could be performed at a lower flow rate (20 nL/min at 150 psi). Although a single pump could perform these operations, the long time required to stabilize the pressure would significantly increase the total analysis time.

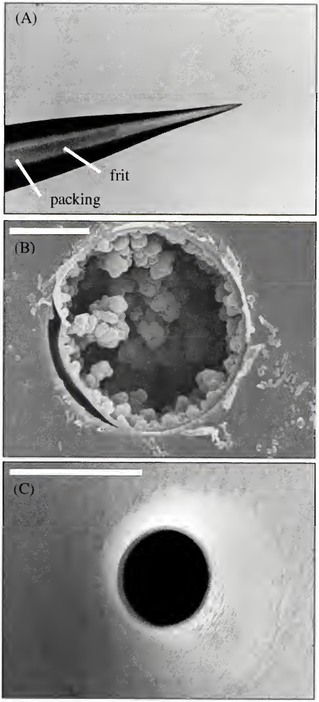


Figure 2-2. CLC columns with integrated electrospray emitters. (A) Bright field optical image of the end of the column with an emitter tip. (B) Scanning electron micrograph (SEM) of macroporous frit formed by *in situ* photo-polymerization (scale bar = 10 μm). After the frit was prepared, the capillary was cleaved to expose the inner column. (C) SEM of electrospray emitter with end-on view (scale bar = 5 μm). Preparation of macroporous frits and CLC columns with integrated electrospray emitters is described in the text.

Gradient Steepness Parameter. In order to achieve good separation and sensitivity, it was necessary to optimize the gradient steepness parameter (b), which is given by equation 2-1:

$$b = \Delta\phi S V_0 / t_G F = 1/1.15k^* \qquad (2\text{-}1)$$

where $\Delta\phi$ is the volume fraction change in organic content during the gradient, S is the solvent strength parameter (the slope of a plot of k' vs. ϕ), V_0 is the column dead volume, t_G is the gradient ramp time, F is the volumetric flow rate, and k* is the average capacity factor[31] (see Appendix A for an example of CLC-MS^2 calculations). It has previously been observed that an increase in b will decrease resolution and increase sensitivity due to decreased band broadening. In other words, various diffusion processes are less pronounced with steeper gradients and shorter separation times, resulting in sharper chromatographic peaks.

In order to investigate these competing effects for the current system, the resolution and sensitivity for Met-enkephalin and Leu-enkephalin were evaluated as a function of b at 70 nL/min as shown in Figure 2-3. The data illustrate the compromise that must be made for sensitivity and resolution with gradient elution. The data also show, however, that decreasing the flow rate to 20 nL/min improved sensitivity and resolution at a given b. (To maintain similar b at lower flow rate, the t_G must be increased as shown in equation 2-1). This improvement may be explained by a combination of effects. The resolution is improved by higher efficiency at lower flow rates (e.g., resolution improves ~ 3-fold at b ~ 3). The sensitivity is improved by the increase in ionization and ion transfer efficiency that is expected with lower flow rates for electrospray ionization[32]. These data illustrate the significant advantage of using low-flow rates for LC-MS with an ESI. Such low-flow rates can be achieved with reasonable b and analysis time if the t_G is appropriate and column i.d. small. The main compromise that was made to maintain similar b values at lower flow rates is the increase in analysis time. Thus, the separation time increased from 1.0 min to 5.0 min for b of 2.0 when going from 70 to 20 nL/min.

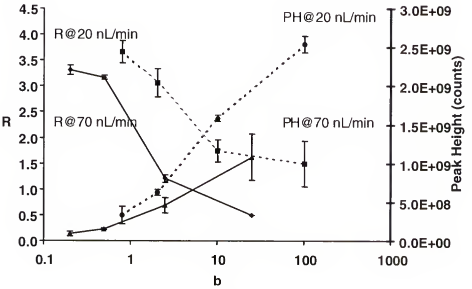


Figure 2-3. Effect of gradient steepness parameter (b) on resolution (R) and sensitivity (peak height) as a function of flow rate. Dashed line is 20 nL/min and solid line is 70 nL/min. R is given by ■ for 20 nL/min and ♦ for 70 nL/min. Peak height is given by ● for 20 nL/min and ▲ for 70 nL/min. Resolution was calculated for 370 nL injections of 18 nM Met- and Leu-enkephalin (6.6 fmol injected on-column). Peak height was determined for Leu-enkephalin. Data are averages from 3 columns (all 2 cm long) and error bars represent 1 standard deviation. The gradient was from 5 to 90% B with $0.1 \le b \le 100$ and $0.1 \le t_G \le 12.5$ min.

For all further experiments, b of 2.0 (20 nL/min flow rate) was used. This b-value is considerably higher than the b of 0.2 that is usually considered a good compromise among resolution, sensitivity, and analysis time[31]. The use of a higher b was dictated by the need to achieve high sensitivity for the trace level detection of neuropeptides while maintaining good chromatographic resolution of Met-enkephalin and Leu-enkephalin.

Previous high sensitivity work on neuropeptides used step gradients (i.e., very high b values), which maximize sensitivity but allow no chromatographic resolution[21]. Furthermore, overly high b values can result in peaks that are too narrow to accurately quantify with the QIT. At the steepest gradient employed in this work (b = 100), the peak width was ~ 0.01 min, which only allows a single scan to be acquired. Employing extremely steep gradients could be useful for qualitative analysis; however, for quantitative analysis it is preferable to collect 8-10 scans across a chromatographic peak. In addition, data-dependent operation of the QIT requires several seconds to complete some scan functions (e.g., full MS, zoom scan, and full MS^2). Therefore, it is desirable to have peaks that are at least a few seconds wide. With the b value and flow rates used here, chromatographic peaks were 0.05 min (3 sec) wide at 0.607 of the peak height (1 σ); that width allowed 8-10 mass spectra to be collected over the width of a peak.

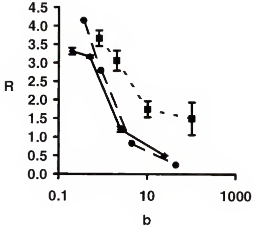


Figure 2-4. Effect of gradient steepness parameter (b) on resolution (R) for different column lengths and flow rates (■ = 2 cm long column at 20 nL/min; ● = 10 cm long column at 50 nL/min; ♦ = 2 cm long column at 70 nL/min). Chromatographic conditions are the same as described for Figure 2-3. Data are the average from 3 columns and error bars are 1 standard deviation.

For these experiments, it was desirable to utilize short columns (2 cm) because a short column will allow faster loading of sample and therefore shorter analysis times, for a given pressure limit; however, it was also desirable to maintain chromatographic separation. Figure 2-4 shows that with manipulation of b and flow rate, it is possible to obtain similar resolution on 2 cm and 10 cm long column. Thus, resolution is similar over a wide range of b for a 10 cm column operated at 50 nL/min and a 2 cm long column operated at 70 nL/min (Figure 2-4). Furthermore, as discussed above, higher resolution can be obtained with lower flow rates (see 20 nL/min data on Figure 2-4). This effect, where resolution depends primarily on b and not L, occurs for molecules with large S-values such as proteins[31]. Our results suggest that, for the peptides used here, S is sufficiently large (~ 3) [31]to achieve a similar independence of L. Therefore, the 2 cm capillary columns do not result in a compromise for separation relative to longer columns.

Figures of Merit. Figure 2-5 illustrates separation with time-segmented MS^2 detection of 61 amol each (1.8 µL of 33 pM) of Met- and Leu-enkephalin injected on-column. The signal-to-RMS noise ratio (S/N) in the reconstructed ion chromatogram (RIC) was 65 (Figure 2-5B), yielding a mass detection limit, based on the amount that gives a S/N of 3, of 4 amol injected on-column, corresponding to a concentration detection limit of 2 pM for a 1.8 µL injection. Linear calibration curves (R^2 greater than 0.995) were obtained from 60 pM to 6 nM. RIC peak height relative standard deviations (RSD) were 20% at 60 pM and improved to 5% at 6 nM. Retention time RSD was 2%. Retention time precision was essential because, if a peak elutes outside of a time-segmented scan function, then it will not be detected. These results are superior to previous trace-analysis work on the QIT where linear calibration curves (R^2 greater than

0.992) were obtained for the drug orlistat in human plasma from 600 pM to 20 nM using an isotopically labeled internal standard[33]. In that work, RIC peak area RSDs were 15% at 600 pM and improved to 6% at 20 nM. While a 2 mm -i.d. column operated at 150 μL/min enabled analysis times of only 3 min, a 20 μL injection volume was required and poor chromatographic resolution was observed with peaks ~ 0.5 min (30 sec) wide[33] (i.e., an ~ 10 fold increase in injection volume and peak width relative to this work). Attomole detection limits have also been reported by Hunt's group; however, no detailed figures of merit are readily available at this time[34-37].

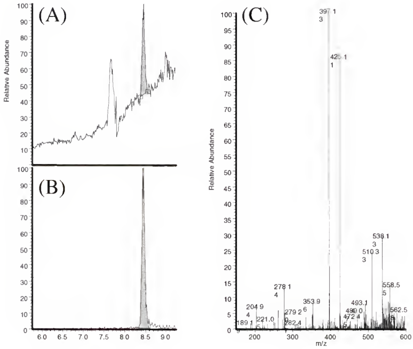


Figure 2-5. High sensitivity CLC-MS^2 measurement of a mixture of Met-enkephalin and Leu-enkephalin at 33 pM (1.8 μL injected corresponding to 59 amol of each peptide loaded onto the column) each in aCSF using a 2 cm column. (A) Time-segmented total ion chromatogram, showing peaks for Met-enkephalin (7.8 min) and Leu-enkephalin (8.5 min). (B) Reconstructed ion-chromatogram for Leu-enkephalin, monitoring the 556→397 + 425 transition. (C) Full scan MS^2 for Leu-enkephalin obtained from the chromatographic peak showing all of the expected b- and y-type ions. The gradient was from 5 to 90% B with b = 2.0, t_G = 5 min, flow rate = 20 nL/min.

Adsorption of peptides within the injection valve (valve carryover) can cause false-positives. This problem is especially prevalent when injecting relatively high concentrations of peptides, required for calibration and tuning, prior to high sensitivity experiments. To avoid this problem, deuterated Leu-enkephalin (YGGF$_{D5}$L) was used as a calibration standard. The effectiveness of this approach is illustrated in Figure 2-6 which shows a RIC from injection of 600 pM YGGF$_{D5}$L (6A), followed by injection of 60 pM YGGFL (6B). (For these experiments, an isolation width = 9 amu was centered at *m/z* 558.5 to fragment YGGFL and YGGF$_{D5}$L simultaneously). The data show that although signal due to carryover for YGGF$_{D5}$L is obtained in the second chromatogram (bottom trace of Figure 2-6B), it does not interfere with the low concentration detection of YGGFL (top trace of Figure 2-6B). Moreover, the a_4 and b_4 ions of YGGF$_{D5}$L at *m/z* 402 and *m/z* 430 , respectively, (6C) are clearly distinguished from the a_4 and b_4 ions of YGGFL at *m/z* 397 and *m/z* 425 (6D).

***In Vivo* Monitoring of Known Endogenous Peptides.** To demonstrate application of the high sensitivity CLC-MS^2 system, dialysate from the rat globus pallidus was analyzed for Met- and Leu-enkephalin. The distribution of the enkephalins in the rat brain has been studied extensively by mRNA *in situ* hybridization[38] and immunohistochemical[39] and immunocytochemical experiments[40]; the globus pallidus is rich with enkephalin-containing neurons. Recently, the distribution of enkephalin-containing neurons in the globus pallidus was found to be heterogenous due to multiple neuronal subpopulations[41,42]. TIC and RIC for Met- and Leu-enkephalin under basal and K$^+$-stimulated conditions are shown in Figure 2-7A and 2-7B respectively. Under basal conditions, no peaks are discernable for the peptides in the TIC (upper trace of Figure 2

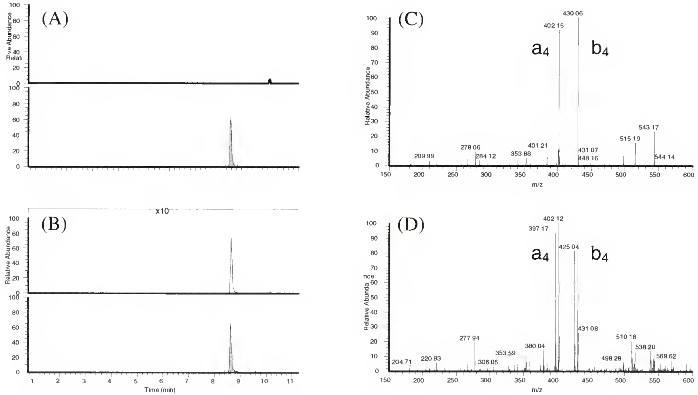


Figure 2-6. Illustration of peptide carry-over in the injection valve and carry-over prevention, using deuterated standards. Reconstructed ion chromatograms from injection of 1.8 μL of 600 pM YGGF$_{D5}$L (A) followed by injection of 1.8 mL of 60 pM YGGFL (B). In (A) and (B) the upper trace is current for the 556→397+425 transition of YGGFL and lower trace is for the *m/z* 561→*m/z* 402 + *m/z* 430 transition of YGGF$_{D5}$L. (C) MS^2 from peak of chromatogram in (A), illustrating the a_4 and b_4 ions of YGGF$_{D5}$L (*m/z* 402 and *m/z* 430). (D) MS^2 from peak of chromatogram in (B), illustrating resolution of a_4 and b_4 ions for YGGF$_{D5}$L and YGGFL (*m/z* 397 and *m/z* 425). Mass spectrometric conditions are described in the text.

7A), although several other peaks are observed. In the time-segmented RIC (lower trace of Figure 2-7A), peaks for the target peptides are readily observed because of the reduced background signal in this mode. During stimulation, the signals for the Met- and Leu-enkephalin dominate the RIC and TIC (Figure 2-7B). The MS^2 mass spectrum acquired for Met- and Leu-enkephalin contains all of the expected b- and y-type ions necessary to sequence Met- and Leu-enkephalin (see Chapter3 for ion assignments) at basal (7C and 7E, respectively) and K^+-stimulated levels (7D and 7F, respectively). Sequence information, combined with retention times that match those of a standard to within 2%, gives confident identification of the detected peaks as the target peptides[26].

Simultaneous monitoring of Met-enkephalin, Leu-enkephalin, and an unknown peptide during a K^+-stimulation is illustrated in Figure 2-8. For monitoring experiments, 2.0 µL samples were collected on-line every 30 min and analyzed (1.8 µL actually injected), resulting in 30 min temporal resolution. The sensitivity of the system would, in principle, allow considerable improvements in temporal resolution if dynamic measurements were required. During a 30 min period, 18 µL of dialysate is generated; however, only 2.0 µL of sample was actually collected for analysis because the dialysate was pumped through the injection loop to waste while one sample was analyzed by LC-MS^2. Therefore, the sensitivity of the system is sufficient to achieve 3.3 min temporal resolution (2.0 µL fractions); however, this would require faster on-line analysis or off-line collection and analysis of fractions.

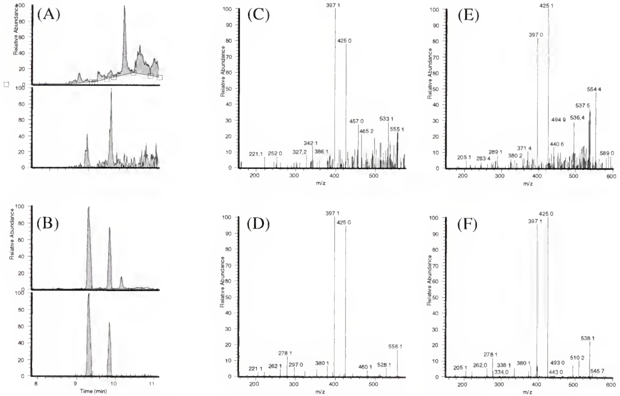


Figure 2-7. *In vivo* detection of Met- and Leu-enkephalin during basal and K^+-stimulated conditions. (A) Total ion chromatogram (top) and reconstructed ion chromatogram (bottom) for a globus pallidus dialysate sample (2.0 μL) collected during basal conditions. The reconstructed ion chromatogram is time-segmented, and shows the current for the 574→ 397 + 425 (8.0 to 9.7 min) and 556→ 397 + 425 (9.6 to 11.0 min) transitions of Met- and Leu-enkephalin, respectively. Met-enkephalin elutes at 9.3 min and Leu-enkephalin at 9.9 min. Chromatographic conditions are described in the text. (C) and (E) show the mass spectra obtained at the elution time of Met-enkephalin and Leu-enkephalin, respectively for chromatogram in (A). (D) and (F) show the mass spectra obtained at the elution time of Met-enkephalin and Leu-enkephalin, respectively for chromatogram in (B).

Basal dialysate levels of Met-enkephalin and Leu-enkephalin were 60 ± 30 pM and 70 ± 20 pM, respectively. The maximum K^+-stimulated levels were 1900 ± 500 pM and 1300 ± 300 pM, respectively (n = 7 rats). As reported previously[2], the levels of Met- and Leu-enkephalin are nearly equal even though their precursor protein, preproenkephalin A (PEA), contains 6 repetitions of the Met-enkephalin sequence and only 1 copy of the Leu-enkephalin sequence. Thus, *in vivo* processing of PEA delivers equal levels to the extracellular fluid despite the presence of more copies of the Met-enkephalin sequence. Previous quantitative work by microdialysis with RIA reported that the total opioid peptide level under basal conditions in an anesthetized rat were approximately 150 pM when correcting for an 8% *in vitro* recovery (Maidment, N.T., personal communication)[18]. Our data give a value of 260 pM when correcting for *in vitro* recovery indicating reasonable agreement. The slightly higher numbers obtained here could be due to several factors including use of an on-line system that prevented loss due to adsorption. In addition, since *in vitro* recovery does not accurately reflect *in vivo* recovery, errors could be introduced in such a comparison because of the difference in dialysis conditions. The K^+-stimulated levels measured for these peptides are about are 2-5 fold higher than those previously reported[2,5,14]. The large difference may be due to use of a more aggressive stimulation (i.e., a 30 min pulse of 150 mM K^+) than previous work (e.g., a 2 min pulse of 50-100 mM K^+)[18], which may increase the amount released, and to the use of the on-line system. The variability of the measured *in vivo* levels (i.e., SEM = 50% for n = 7) was greater than that seen for *in vitro* measurements (i.e., peptide standards dissolved in aCSF) at the same concentration (i.e., SEM = 12% for n = 3). This suggests that a large part of the irreproducibility is due to variables in the *in vivo* measurement such as probe positioning or inter-animal differences.

***In Vivo* Identification of Unknown Endogenous Peptides.** During time-segmented MS^2 experiments, unknown compounds were observed that were co-released with Met- and Leu-enkephalin. These compounds could be monitored along with the enkephalins, as illustrated by the example in Figure 2-8 (see also the peak at 10.1 min in the top trace of Figure 2-7B). The detection of these compounds even with a narrow isolation width (3 amu) employed for time-segmented MS^2 suggested that it might be possible to identify unknowns collected from the extracellular space. To explore this possibility, data-dependent MS^2 analysis of dialysates was performed. A relatively narrow scan range in the MS mode (*m/z* 550 - 600) was employed during data-dependent MS^2 experiments in order to maintain high sensitivity. Data-dependent MS^2 scans of a sample collected during K^+ stimulation (Figure 2-9A) combined with time-segmented MS^2 data revealed 11 novel peptides. In addition to these unknowns, high quality MS^2 spectra for the known peptides Met- and Leu-enkephalin were also obtained in this mode (Figure 2-9C and 2-9D). The data-dependent MS^2 spectrum for one of the unknown peptides, which eluted at 7.2 min, is shown in 9B. The fragmentation pattern for this unknown suggested a doubly charged peptide ion; however, only a partial sequence was obtained because of the product-ion scan range used. Subsequent high-resolution zoom scans (data not shown) confirmed that the unknown peptide was indeed doubly charged. Database searching (Sequest) identified the peptide as SPQLEDEAKE, a sequence found only in PEA. To confirm the identification of this peptide, time-segmented MS^2 and MS^3 scans were performed and compared to synthetic peptide as shown in Figures 10 and 11. The retention time (within 2%) and fragmentation pattern (all of the expected b- and y-type ions were observed) were in excellent agreement with the synthetic peptide (see Chapter3 for ion assignments). In addition, the ion intensities of the standard were within 30% of those for the *in vivo* sample. Combined, this information gives confident sequence assignment.

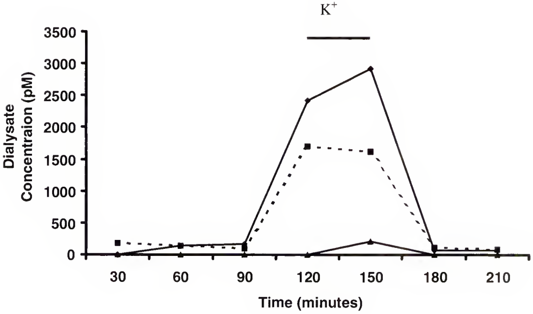


Figure 2-8. ***In vivo*** monitoring results for Met-enkephalin (■), Leu-enkephalin(●), and for an unknown peptide (▲) for a single rat. Each data point is from a 2 μL sample collected at the time indicated. The bar indicates application of 150 mM K^+ to the microdialysis probe.

The TIC and RIC for the novel peptide in MS^2 mode (see lower trace of Figure 2-10A) shows a surprising amount of chemical noise. This noise may reflect the complexity of the dialysate samples. Interestingly however, the S/N is considerably improved for the MS^3 TIC (Figure 2-11A), suggesting that for some applications this mode of detection may provide improved sensitivity, due to reduced background, over MS^2 detection. However, this possibility was not explored.

From *in vivo* samples, the SPQLEDEAKE was never present above 400 pM (800 amol injected on-column and 2-3-fold less than Met- and Leu-enkephalin), yet the system was able to identify this peptide from the dialysate. In principle, this approach could be used over a wider scan range to identify more peptides in the dialysate; however, with the weak signals observed (all peptides are at the amol level), it may be best to perform such characterizations using multiple injections with overlapping, narrow (*m/z* 50) MS scan ranges (see Chapter3). Although 11 peptides were detected and tentatively identified in these data-dependent runs, further analysis, as performed for SPQLEDEAKE, is required to confirm identification of these peptides.

PEA Processing. As indicated above, SPQLEDEAKE is a peptide sequence found in PEA. Extensive prior studies have not revealed the presence of this peptide as a product of PEA processing[43,44,44,45,45]. Because most work on peptide processing uses antibodies to select peptides with certain epitopes, and because this peptide does not contain any homology with Met- or Leu-enkephalin, it is unlikely that this peptide would have been detected by any conventional processing studies. As depicted in Figure 2-12, this peptide is an N-terminal cleavage product of peptide I (peptide I_{1-10}). Peptide I_{1-10} does not appear to have been produced from peptide I by the cleavage at dibasic sites that is typically associated with endopeptidase activity. The N-terminus appears to have been produced by cleavage at dibasic sites; the C-terminus is not. A reasonable hypothesis for the production of peptide I_{1-10} is endopeptidase cleavage of peptide I, a known product of PEA in purified bovine chromaffin cells[46], to yield SPQLEDEAKELQ which is then cleaved by a carboxypeptidase to yield the SPQLEDEAKE sequence detected *in vivo*. This pathway seems likely, given the dibasic site in peptide I and the apparent

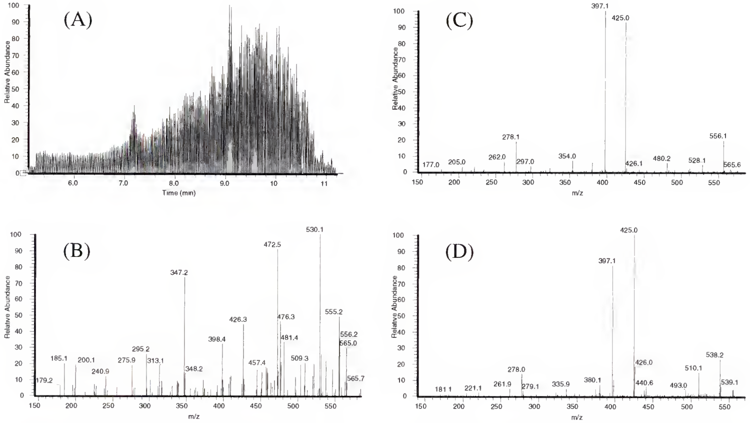


Figure 2-9. *In vivo* identification of peptides by data-dependent LC-MS^2. (A) Total ion chromatogram for data-dependent MS^2 of a dialysate sample collected from the globus pallidus during K$^+$-stimulation. (B), (C), and (D) show MS^2 of SPQLEDEAKE, Met-enkephalin, and Leu-enkephalin, respectively.

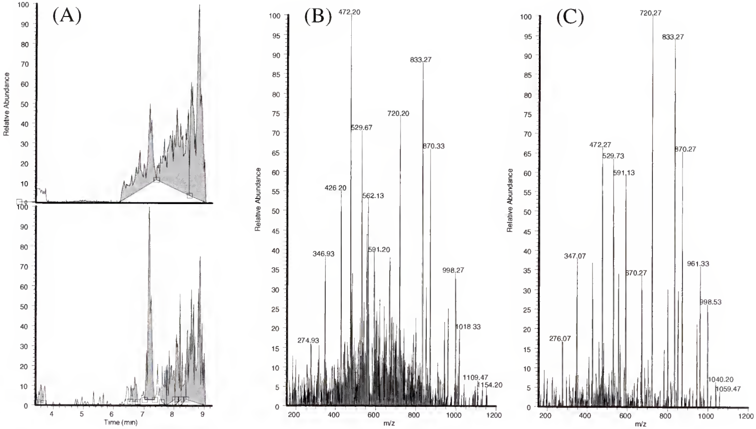


Figure 2-10. Sequencing of a novel peptide by time-segmented MS^2 from sample collected *in vivo* during K^+-induced depolarization. (A) Total ion chromatogram (top) and reconstructed ion chromatogram (bottom) for the 574→ 472 + 529 transition. (B) and (C) show MS^2 of SPQLEDEAKE obtained *in vivo* and from synthetic peptide, respectively. Synthetic peptide was injected at 600 pM.

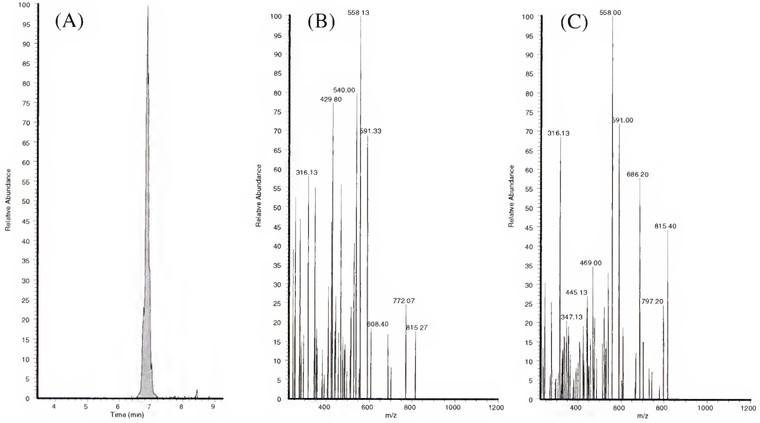


Figure 2-11. Confirmation of the sequence of a novel peptide by time-segmented MS^3 from sample collected *in vivo* during K$^+$-induced depolarization. (A) shows the total ion chromatogram for the 574→833→ transition. (B) and (C) show MS^3 of SPQLEDEAKE obtained *in vivo* and from synthetic peptide (600 pM), respectively.

ubiquity of exopeptidase activity in the brain extracellular space[47]. The physiological role of peptide $I_{(1\text{-}10)}$ (if any) and confirmation of the mechanism for its production from PEA remains to be determined.

MAQFLRLCIWLLALGSCLLATVQADCSQDCAKCSYRLVRPGDINFLACTLECEGQLPSFKIWETCKDLL
QVSKPEFPWDNIDMYKDSSKQDESHLLAKKYGGFMKRYGGFMKKMDELYPVEPEEEANGGEILAKR
YGGFMKKDADEGDTLANSSDLLKELLGTGDNRAKDSHQQESTNNDEDSTSKRYGGFMRGLKR**SPQL**
EDEAKELQKRYGGFMRRVGRPEWWMDYQKRYGGFLKRFAESLPSDEEGESYSKEVPEMEKRYG
GFMRF

Figure 2-12. Amino acid sequence of PEA. Sequences matching Met-enkephalin and Leu-enkephalin are underlined; however, these peptides are not necessarily released by PEA processing. Peptide I, which has been observed in previous work on tissues[26] is in bold. The novel peptide found in this work is shown in the box.

Conclusion

Twenty-five μm inner diameter (i.d.) fused silica CLC columns with 3 μm i.d. integrated electrospray emitters interfaced to a QIT mass spectrometer were evaluated for high sensitivity CLC-MS^2. Column preparation involved constructing frits by *in-situ* photo-polymerization of glycidyl methacrylate and trimethylolpropane trimethacrylate, preparing the electrospray emitter by pulling the column outlet to a fine tip with a CO_2-laser puller, and slurry-packing the column with 5 μm reversed-phase particles. Large volume injections were facilitated by an automated two-pump system that allowed high-flow rates for sample loading and low-flow rates for elution. Small electrospray emitters, low elution flow rates and optimization of gradient steepness allowed a detection limit of 4 amol, corresponding to 2 pM for 1.8 μL injected on-column, for a mixture of peptides dissolved in artificial cerebral spinal fluid. The system was coupled on-line to microdialysis sampling, and was used to monitor and discover endogenous neuropeptides

from the globus pallidus of anesthetized male Sprague-Dawley rats. Time-segmented MS^2 scans enabled simultaneous monitoring of Met-enkephalin, Leu-enkephalin and unknown peptides with at least 30 min temporal resolution. Basal dialysate levels of Met-enkephalin and Leu-enkephalin were 60 ± 30 pM and 70 ± 20 pM while K^+-stimulated levels were 1900 ± 500 pM and 1300 ± 300 pM, respectively (n = 7). Data-dependent and time-segmented MS^2 scans revealed several unknown peptides that were present in dialysate. One of the unknowns was identified using MS^2, MS^3, and database searching as peptide I_{1-10} (SPQLEDEAKE), a novel product of PEA processing. This method should be useful for monitoring known peptides, studying processing of endogenous peptides and characterizing novel peptides that are potentially neuroactive.

CHAPTER 3
DISCOVERING ENDOGENOUS NEUROPEPTIDES BY *IN VIVO* MICRODIALYSIS-CLC-DATA-DEPENDENT-MS[2]

Introduction

Neuropeptides are inter-neuron signaling molecules important in many physiological and behavioral functions including learning, memory, appetite regulation, sleep, sensory perception, immunity and disease. Neurons produce neuropeptides by synthesizing protein precursors which are packaged into vesicles for storage and secretion. Proteases cleave precursors within the vesicles to form a collection of peptides that are released by exocytosis when the neurons are depolarized. After secretion from neurons, the peptides can be further processed by peptidases in the extracellular space. Such processing can both activate and deactivate secreted peptides (e.g., cleavage of various enkephalin-containing peptides produces the enkephalins and many peptides with no known bioactivity). It has recently been shown that proteolytic processing produces neuropeptides that are not predicted from known protease cleavage sites[47-51] (see Chapter 4). Thus, the complete mechanisms regulating proteolytic processing remain unknown. Besides neuropeptides, fragments of other proteins may be expected to be in the extracellular space and have biological significance. For example, extracellular accumulation of β-amyloid pepide, derived from the amyloid precursor protein, is observed in Alzheimer's disease[52].

One route to study processing of neuropeptides is by analysis of brain tissue. The advent of proteomic techniques has greatly expanded the power of this approach. In this method, peptides formed by tissue proteases are isolated from the sample and then analyzed by capillary liquid chromatography (CLC)-tandem mass spectrometry (MS^2). The resulting MS^2 spectra can be correlated to protein precursors by database searching. This method is similar to the "shot-gun" proteomics method in which a protein mixture is pretreated by selective proteolysis (e.g., trypsin digestion) to form a collection of peptides characteristic of the proteins and the protease used for proteolysis prior to CLC-MS^2 analysis and database searching. Application of this method to peptides in tissue is complicated by the lack of control over the proteases used for digestion, i.e.,not all of the peptides are tryptic peptides. This creates at least two difficulties. Tryptic peptides have basic sites resulting in good ionization efficiencies and efficient cleavage along the peptide backbone promoting a high yield of b- and y-type ions[16]. Peptides produced by other proteases may be less sensitively detected. In addition, without *a priori* knowledge of the protease specificity, a greater library of peptides must be searched thus increasing the probability of finding a random match. Despite these limitations, many new peptides have been found by this approach[49-51].

A problem with tissue analysis for studying neuropeptides is that no discrimination is made between intracellular peptides and those that are actually released into the extracellular space. In addition, tissue analysis could not be used for correlating behavior to a given set of released peptides. Analysis of peptides in the extracellular space of live animals is required for such studies. Recently peptides in the extracellular compartment of live rats have been detected by microdialysis sampling coupled with

CLC-MS[2] or CE-MS[53]; however, these studies have relied on detecting just a few known peptides or infusing a known peptide at high levels to follow peptide processing *in vivo*. Such approaches preclude identification of novel endogenous species.

The objective of this work was to develop and evaluate a method for exploring the endogenous peptide content and proteome of brain extracellular fluid in living animals. Development of such a method would allow many novel studies such as identification of processing patterns after release, correlation of peptides released into the extracellular fluid with different behavioral and physiological states, and identification of potentially novel neurotransmitters. Extremely low concentrations of endogenous peptides in dialysate (1-100 pM) and small sample volumes (1-100 μL) available place stringent sensitivity requirements on the analytical methods used[18,21,22,24,54]. In previous work, we demonstrated that CLC columns with integrated electrospray emitters operated at low nanoliter/min flow rates interfaced to quadrupole ion traps could attain attomole detection limits for targeted peptides in microdialysis samples. That prior work also demonstrated identification of a novel endogenous peptide while operating the mass spectrometer over a narrow mass-to-charge ratio (*m/z*) range in the MS mode (*m/z* 50). In the present work, we demonstrate the use of a much wider m/z range in the MS mode (*m/z* 1500) and confirm identification of 29 peptides (27 novel) as well as preliminary identification of over 200 other peptides.

Experimental

Chemicals and Reagents. CLC solvents were purchased from Burdick and Jackson (Muskegon, MI). Peptides were from Sigma (St. Louis, MO). Acetic acid and hydrofluoric acid were purchased from Fisher Scientific (Pittsburgh, PA). Ethanol was purchased from J.T. Baker (Phillispburg, NJ). The following chemicals were purchased from Aldrich (Milwaukee, WI) to prepare macroporous photopolymer frits: isooctane, toluene, trimethylolpropane trimethacrylate, glycidyl methacrylate, and benzoin methyl ether. Artificial cerebral spinal fluid (aCSF) used for microdialysis perfusion consisted of 145 mM NaCl, 2.68 mM KCl, 1.10 mM $MgSO_4$, and 1.22 mM $CaCl_2$. The high K^+, low Na^+ perfusate solution for depolarization experiments consisted of 145 mM KCl and 2.62 mM NaCl with other salts the same as aCSF. Mobile phases and aCSF were prepared weekly, and were filtered with 20 nm-pore size aluminum oxide filters (Fisher Scientific) to remove particulates. Tryptic digests were purchased from Michrom Bioresources (Auburn, CA).

***In Vivo* Microdialysis.** Commercial microdialysis probes (CMA/10 with 20 kDa molecular weight cut-off, CMA, Solna, Sweden) with 4-mm active lengths were used as described previously by radio-immunoassay[6,18], tandem mass spectrometry[23,24], and CLC with electrochemical detection[21,22]. The outlet tubing of the microdialysis probe was replaced with fused silica capillary (50 μm i.d., 360 μm o.d., Polymicro Technologies, Phoenix, AZ) to minimize dead volumes. Prior to use, microdialysis probes were rinsed with 70% ethanol at a flow rate of 0.6 μL/min per the manufacturer's instructions. Deuterated Leu-enkephalin (University of Florida Protein Core), where 5 non-exchangeable deuterium atoms were incorporated into the phenyl ring of phenylalanine, was used to calculate the relative recovery through the microdialysis probe and confirm sensitivity prior to *in vivo* experiments. *In vitro* relative recovery of 600 pM deuterated Leu-enkephalin in aCSF was 51 ± 16% (n = 4) at 37 °C. Surgical preparation and procedures were the same as described previously[21]. In this work, the microdialysis probe was implanted into the striatum region of the rat brain with the following coordinates relative to the bregma region: anterior-posterior = -1.0 cm, medial-lateral = +0.17 cm and dorsal-ventral = -0.91 cm. Correct placement of the microdialysis probe into the striatum was confirmed by inspection of cresyl violet-stained tissue sections cut by a cryomicrotome.

Preparation of CLC Columns with Integrated Electrospray Emitters. To prepare columns, a 1 cm section of polyimide was removed from a 20 cm length of 25 μm i.d. fused silica with a flame. The capillary was loaded with a solution of benzoin methyl ether (1.5% w/w), glycidyl methacrylate (7% v/v) and trimethylolpropane trimethacrylate (16% v/v) in a 3:1 isooctane:toluene solvent, using He pressure, and was placed inside PEEK sleeves (Upchurch Scientific, Oak Harbor, WA) that formed a mask that exposed a ~100 μm length of the bare fused silica. Polymerization was achieved by illuminating the exposed region of the capillary with a UV lamp (Spectronics, Westbury, NY) for 30 min[28,29]. Frits were flushed with acetone and dried for storage by passing He through the capillaries for 5 min at 500 psi. A P-2000 CO_2 laser puller (Sutter Instruments, Novato, CA) was used to create integrated electrospray emitters less than 1 cm beyond the frit. The settings used on the P-2000 were heat 350, filament 0, velocity 20, delay 135, and pull 125. The program was cycled 3-4 times to generate pulled capillary columns with less than 1 μm orifices. The resulting pulled columns were etched with 50% hydrofluoric acid (use extreme caution and neutralize with $CaCl_2$)[30] for 30 s to create sharp-edged electrospray emitters with approximately 3 μm i.d. Pulled columns were packed with an acetone slurry (10 mg/mL) of 5 μm Alltima C18 reversed-phase particles (Alltech, Deerfield, IL) at 1000 psi as described elsewhere.[18] Only 2 cm of the total 13 cm length of fused silica was packed, unless stated otherwise. The void at the head of the column did not contribute to extra-column band-broadening because all samples were injected in weak mobile phases to allow the analytes to stack at the head of the column.

Automated Two-Pressure CLC-MS^2 System and Operation. The CLC-MS^2 system and its operation are described in detail elsewhere[6]. The system utilizes two six-port valves to select the pump and flow path for preconcentration, desalting and separation/electrospray steps. During the preconcentration and desalting steps the high flow rate pump is selected without splitting of the sample in order to minimize the sample loading time. During the separation/electrospray step the low flow rate pump is selected with splitting of the gradient in order to maximize the separation and electrospray efficiency and to minimize the delay time of the gradient, respectively.

In this work, 8.3 min microdialysis fractions, collected online in a 5 μL sample loop, were analyzed every 35 min by preconcentrating/desalting at 370 nL/min and separating/electrospraying at 10 nL/min. All measurements were made with the following CLC-MS^2 parameters, unless specified otherwise: preconcentration time = 10 min (3.6 μL), desalting time = 5 min (1.8 μL), separation/electrospray time = 15 min, re-equilibration time = 5 min. The mass spectrometer was a quadrupole ion trap (LCQ-Deca, ThermoFinnigan, San Jose, CA) with the following parameters, unless specified otherwise: automatic gain control (AGC) on, max AGC time = 300 ms, q CID = 0.25, isolation width = 3 amu normalized collision energy = 35%, activation time = 0.25 ms and the default number of microscans and target count values. Data-dependent MS^2 spectra were collected in the 'triple play' scan mode (MS, zoom, MS^2) for precursor ions with a total ion current greater than 1 x 10^5 counts using precursor ion windows of *m/z* 550-600 and *m/z* 500-2000 for 3 and 10 rats, respectively. Dynamic exclusion (exluding precursor ions from repeated analysis) was enabled in order to minimize the collection of redundant spectra, extend dynamic range and characterize samples more comprehensively..

Spectral Reproducibility. Spectral reproducibility describes the percentage of the MS^2 spectra within any two CLC-MS^2 datasets (n and n+1) that are the same for a given precursor ion mass tolerance (e.g., 2.5 Da) and product ion threshold (e.g., 38% correlation). Spectral reproducibility was determined by the following formula:

$$\text{Spectral Reproducibility} = 100\ (n - [(n + 1) - n] / n)$$

where n is the number of MS^2 spectra for the 1st dataset, (n + 1) is the number of MS^2 spectra for the 2nd dataset and [(n + 1) - n] is the number of MS^2 spectra remaining following subtractive analysis of the two datasets.

Subtractive Analysis. Subtractive analysis is a method whereby the difference MS^2 spectra between any two CLC-MS^2 datasets can be determined for a given precursor ion mass tolerance and product ion threshold. Subtractive analysis was performed by the Sequest program Ionquest[55] (a binning algorithm). This algorithm is usually used to subtract known contaminants (e.g., trypsin autolysis products and plasticizers) from CLC-MS^2 datasets prior to peptide sequencing and protein precursor identification by Sequest. In this work, a precursor ion mass tolerance of 2.5 Da and a product ion threshold of 38% correlation (i.e., the default values) were used to 1) determine spectral reproducibility and 2) determine MS^2 spectra observed only during depolarization.

Peptide Sequencing and Protein Precursor Identification. The Sequest Xcorr (cross correlation) score is a measure of the similarity between the *m/z* of the product ions observed in the MS^2 spectrum and the product ions predicted for peptide sequences generated by *in silico* proteolysis of the protein database. The normalized difference between the 1st-ranked and 2nd-ranked peptide sequences is the Sequest ΔCn (delta normalized correlation) score. Xcorr scores greater than 2.0 and ΔCn scores greater than 0.1 are considered significant[56]. However, it was recently shown that Sequest Xcorr scores below 2.5 for doubly-charged precursor ions and below 3.0 for triply-charged precursor ions suggest that false-positives will be present[57,58].

In order to minimize false-positives, a six step data-reduction strategy incorporating database searching, automated *de novo* sequencing and library searching was investigated for peptide sequencing and protein precursor identification. The data-reduction strategy *in vivo* and definitions characterizing the MS^2 spectra remaining after each step of data-reduction (1-6) are summarized below. The strategy involves sequential application of criteria, each set of which provides increasing selectivity.

Six Step Data-Reduction Strategy.

1. MS^2 spectra remaining, from those MS^2 spectra collected, with a total ion current greater than 5×10^5 counts.
2. MS^2 spectra remaining from step 1 following subtraction of MS^2 spectra collected under basal conditions (with artificial CSF perfusing the probe) from MS^2 spectra collected during localized depolarization evoked by infusion of high K^+, low Na^+ solution through the dialysis probe. These MS^2 spectra were considered depolarization-specific.

  a. Subtractive analysis was performed by the Sequest program Ionquest[55] (a binning algorithm) as described above.
  b. Database searching programs (Sequest[55] ver 1.2 and Mascot[59] ver. 1.8) correlated uninterpreted MS^2 spectra with the NCBI nonredundant protein database (i.e., identical sequences are merged into one entry) for *Rattus norvegicus* (12,670 sequences). Precursor ion and product ion mass tolerances were 1.4 and 0.0 for Sequest and 2.0 and 0.8 for Mascot, respectively, as suggested by the manufacturers (the programs have different scoring systems[1]). No protease specificity was selected. Post-translational modifications (PTMs) were investigated by performing error tolerant searches with Mascot (see www.matrixscience.com for a complete list of the hundreds of PTMs searched with error tolerant settings).

3. MS^2 spectra remaining from step 2 with a Sequest ΔCn score greater than 0.1 were considered significant.
4. MS^2 spectra remaining from step 3 that met the following criteria were considered validated by database searching.

---

[1] Important scores used in Sequest (Mascot's scoring system is not publicly available).

- Sp = (Σim)ni(1+β)(1+q)/nT

where im is the matched intensities, ni is the number of matched fragment ions, β is the type b- and y-ion continuity, q is the presence of immonium ions and their respective amino acids in the predicted sequence and nT is the total number of predicted fragment ions

- Cn =normalized integral(x(t)*y(t+tau)dt where T is displacement value
- Xcorr = integral(x(t)*y(t+tau)dt where T is displacement value
- ΔCn = the difference between the 1st-ranked Xcorr and the 2nd-ranked Xcorr

   a. The 1st ranked Sequest-derived sequence matched the 1st ranked Mascot-derived sequence with a Sequest ΔCn score greater than 0.1 and a Mascot score indicating homology with greater than 95% probability (the rigorous definition of this is proprietary).
   b. More than one peptide was observed from each protein precursor.

5. MS^2 spectra remaining from step 4 were evaluated by automated *de novo* sequencing using the Lutefisk program[60] (ver. 1.3.2). A precursor ion tolerance of 2.0 and a product ion tolerance of 0.8 were employed (consistent with Mascot). If the 1st ranked *de novo*–derived partial sequence matched the 1st ranked database-derived sequence with greater homology than it matched the 2nd ranked database-derived sequence peptide sequence, then the MS^2 spectrum was considered validated by *de novo* sequencing-as previously described[60].
6. MS^2 remaining from step 3 were examined for incorrect assignments by the database searching programs. This done by searching the MS^2 spectra from step 4 against a library of MS^2 spectra made from step 3. Manual inspection of the difference spectrum (subtraction of matching MS^2 spectra regardless of precursor ion mass or product ion threshold) was used to confirm a high degree of similarity between matching MS^2 spectra. Library searching was performed using Xcalibur (ver. 1.2) in conjunction with NIST MS Search (ver. 1.7). A simple similarity search was selected with the default NIST MS Search options.

Results and Discussion

Our experiments had two primary goals. The first was to develop and assess an approach to evaluate the peptide and proteome content of the brain extracellular space.

The second was to evaluate the results obtained from initial application of this approach to K+-stimulated release of peptides in the striatum of anesthetized rats.

Spectral Reproducibility. Protein identification by shot-gun proteomics is a well-established method with a high success rate[15]. Application of this tool to endogenous peptides in dialysis samples is complicated by: 1) the need to confirm the peptide sequence, not just the protein identity, 2) unknown protease specificity in peptide formation and 3) limited sample availability and low concentrations[61]. Therefore, prior to *in vivo* experiments, we explored the reproducibility of peptide sequencing and protein precursor identification at the levels expected in our experiments. In addition, we hypothesized that spectral reproducibility would be important in order to correlate MS^2 spectra via database and library searching. Unassigned and incorrectly assigned MS^2 spectra are possible due product ion signals of low signal-to-noise (S/N) for less abundantly released peptides and non-peptide contaminants (e.g., plasticizers). Unsuccessful assignment can be a problem even in relatively simple proteomics experiments; for example, 60% of the MS^2 spectra observed from a single gel slice remained unassigned following database searching[62].

We hypothesized that searching a small user library, in addition to a large protein database, would reveal incorrect sequence assignments at attomole levels. It is well known that mass spectral libraries function well only if good spectral reproducibility is achieved[55,63,64]. Therefore, we investigated the impact of spectral reproducibility on peptide sequencing and protein precursor identification at attomole levels. We also constructed a library of MS^2 spectra in order to find incorrect assignments by the database searching programs.

Peptide Sequencing and Protein Precursor Identification *In Vitro*. Reproducibility was initially assessed by injecting 0.4 µL of a mixture containing a 1 nM tryptic digest of three proteins (lysozyme, carbonic anyhdrase II, and conalbumin) spiked with 1 nM $YGGF_{D5}L$ in aCSF for CLC-MS^2 analysis with data-dependent scanning using a *m/z* 500-2000 precursor ion window. (The concentration of each tryptic peptide is ~ 1 nM). These analyses yielded 132 ± 22 (n = 3) data-dependent MS^2 spectra per chromatogram with a *m/z* 500-2000 precursor ion window. All errors are reported at the 95% confidence interval for comparison with *in vivo* results[2]. Examination of the MS^2 reconstructed ion chromatograms (MS^2 RIC) in Figure 3-1 A-F showed good reproducibility for the *in vitro* samples. This observation was supported by subtractive analysis, which indicated that 75 ± 18% of the MS^2 spectra within any two datasets were the same for these *in vitro* samples. The trace level peptide $YGGF_{D5}L$ (400 amol injected on-column) was observed in every case with good spectral reproducibility (Figure 3-1 G-I).

We were also interested in determining the reproducibility of peptide sequencing and protein precursor identification. Database searching revealed that 62 of the 395 MS^2 spectra collected (16 ± 7% or 21 ± 6 spectra per sample) were significant (i.e., achieved a Sequest ΔCn score greater than 0.1). From these, 40% or 25 MS^2 spectra (8 were unique) were validated by identifying the same peptide sequence and protein precursor by Sequest and Mascot with at least two peptides per protein precursor (see experimental section). Seven of the 8 uniquely validated MS^2 spectra were observed in all of the samples. All three protein precursors were successfully identified with 3 to 29% sequence coverage. Sixty percent or 37 of the 62 significant MS^2 spectra were not validated by a Mascot match, 60% or 37 of 62 were not validated by at least two peptides per protein precursor (i.e., no additional MS^2 spectra were eliminated by requiring at least two peptides per protein precursor) and 27% or 17 of 62 were redundant (i.e., observed more than once.

---

[2] The error in the measurements is described by the confidence interval (μ) and not the standard error of the mean (SEM) because it enables a more correct comparison between datasets. The SEM is the standard deviation (s) of the sampling distribution ($n^{0.5}$) of the mean (x). This formula does not assume a normal distribution, and the SEM simply shows that s is inversely proportional to the square root of the number of observations (n). Significantly, μ gives us the uncertainty of the SEM from a limited n. This uncertainty is expressed by the width of μ at a specified probability (e.g., 95%). These terms are related by the expression [$\mu = x \pm ts/n^{0.5} = x \pm t(SEM)$], where x is the sample mean and t is the

We also evaluated the suitability of *de novo* sequencing to further validate the MS^2 spectra. Seventy-five percent (6 of 8 peptide sequences) of the MS^2 spectra validated by database searching were also validated by *de novo* sequencing by showing that the 1st ranked *de novo*–derived partial sequence (Lutefisk program) matched the 1st ranked database-derived sequence better than it matched the 2nd ranked database-derived sequence[60]. In the case of peptide VGDANPALQK from carbonic anhydrase II observed in sample 2, the MS^2 spectrum yielded a Sequest ΔCn equal to 0.31 and a Mascot score indicating homology with greater than 95% probability. The 1st ranked *de novo*–derived partial sequence from Lutefisk was VGDAN[168.1]LKK, while the 2nd ranked sequences from Sequest and Mascot were GTGKLVALKK and TVAGQVLAKK, respectively.

Student's t-value for a specified probability and number of degrees of freedom (e.g., t = 2.228 for 10 degrees of freedom, n = 11, at the 95% confidence interval).

De novo sequencing was unable to validate database-derived sequences for 2 of the 8 peptides validated by database searching. For example, the peptide FESNFNTQATNR from lysozyme yielded a Sequest ΔCn equal to 0.56 and a Mascot score indicating at least homology with 95% probability. The 1st ranked *de novo*–derived partial sequence from Lutefisk was [276.1]LAKTNFAM[257.1], while the 2nd ranked sequences from Sequest and Mascot were EFAEYVTNHYR and YEPAQIHLSNTR, respectively. In this case, no sequence homology was observed between the database-derived results and the *de novo* sequencing-derived results (despite a high quality MS^2 spectrum with Sequest Xcorr = 3.7). From these results we conclude that validation by *de novo* sequencing is confirmatory; however lack of a match between *de novo* sequencing and database searching does not invalidate the sequence identification made by database searching. This is in excellent agreement with results reported previously where the 1st ranked database-derived sequence was found to be more accurate than the 1st ranked *de novo*-derived sequence[60].

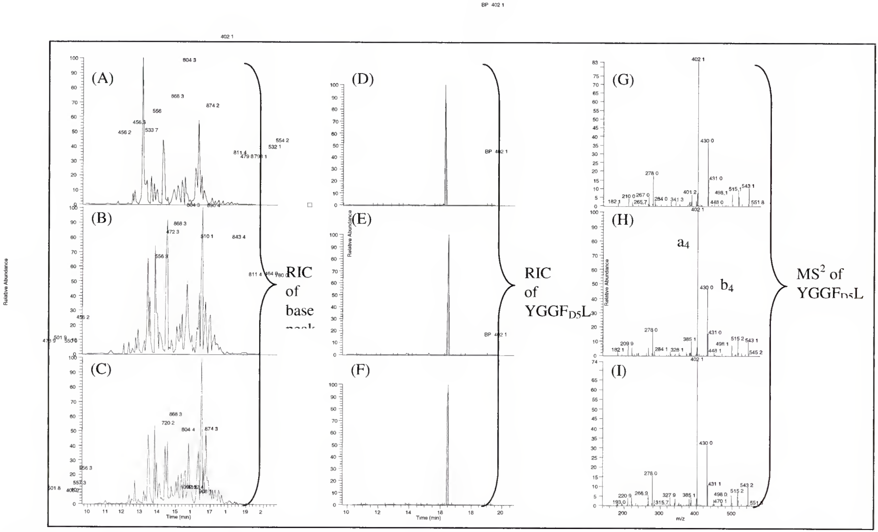


Figure 3-1. Triplcate, 0.4 µL, injections of a tryptic digest of lysozyme, carbonic anyhdrase, and conalbumin at 1 nM with 1nM YGGFD$_5$L (400 amol injected on-column). Base peak reconstructed ion chromatograms (RIC) show the most abundant precursor ions observed *in vitro* (A-C) and the product ions for the YGGF$_{D5}$L precursor ion at *m/z* 561 (D-F). Data-dependent MS^2 spectra of the precursor ion at *m/z* 561 in (G-I) show the characteristic a$_4$ and b$_4$ product ions (*m/z* 402 and *m/z* 430, respectively) of YGGF$_{D5}$L. Peaks are labeled by the most abundant product ion of the data-dependent MS^2 spectrum (e.g., *m/z* 402 for YGGF$_{D5}$L).

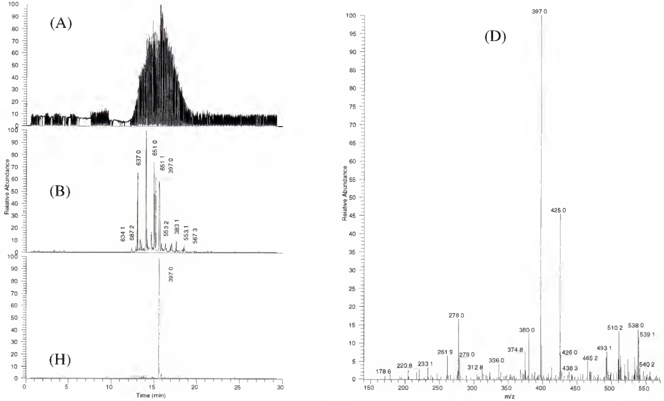


Figure 3-2. *In vivo* microdialysis-CLC-MS^2 analysis of CSF collected from the rat striatum during K$^+$-induced release conditions. (A) Total ion chromatogram, (B) base peak RIC for the most abundant product ion and (C) RIC for the *m/z* 556 → *m/z* 397 + *m/z* 425 transition of YGGFL. Peaks are labeled by the most abundant product ion of the data-dependent MS^2 spectrum (B and C) (e.g., *m/z* 397 for YGGFL). (D) Data-dependent MS^2 spectrum of the precursor ion at *m/z* 556 in (B) shows the characteristic a_4 and b_4 product ions (*m/z* 397 and *m/z* 425, respectively) of YGGFL.

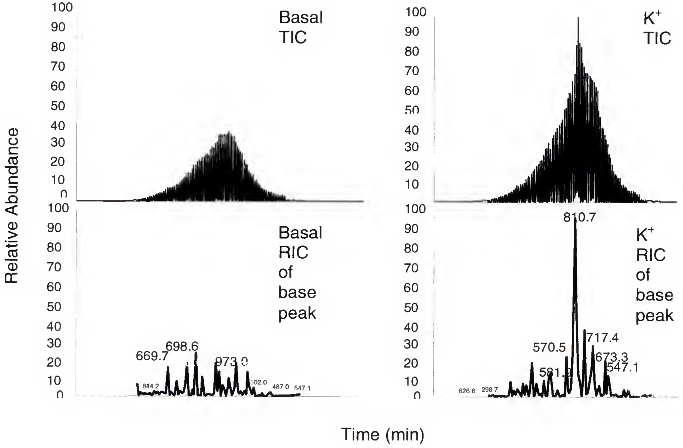


Figure 3-3. Comparison of total ion chromatograms (TICs) and reconstructed ion chromatograms (RICs) for data-dependent MS^2 of dialysate collected from rat striatum during basal (B and D, respectively) and depolarizing (A and C, respectively) conditions showing the most abundant precursor ions observed *in vivo*. Peaks are labeled by retention time in total ion chromatograms (A and B) and by the most abundant product ion of the MS^2 spectrum in reconstructed ion chromatograms (C and D).

Table 3-1. Summary of *in vivo* microdialysis-CLC-MS^2 results.

Animal (n)	Spectral Reproducibility			Data-Reduction Step													
				1			2			3		4		5		6	
	Animals	$K^{+}_{n} - K^{+}_{(n+1)}$	*% Match*	*Basal*	K^{+}	*Total*	$K^{+}_{n} - \text{Basal}_{n}$	*% Match*	*% Total*	*No.*	*% Total*	*No*	*% Total*	*No*	*% Total*	*No*	*% Total*
1	1-2	39	73	149	144	293	51	65	17	34	12	1,4,7,8 (4)	1	3	1	4(1)	0.3
2	2-3	83	33	128	123	251	56	54	22	22	9	7,8 (2)	1	1	0	0	0.0
3	3-4	94	34	148	142	290	49	65	17	20	7	none	0	0	0	0	0.0
4	4-5	80	16	22	95	117	74	22	63	21	18	4,6,8,14 (4)	3	4	3	0	0.0
5	5-6	109	24	138	144	282	85	41	30	28	10	9,11,18,21-25,27 (9)	4	6	2	9,11,16,17 (4)	1.4
6	6-7	93	29	70	131	201	90	31	45	27	13	10,11,28 (3)	1	1	0	0	0.0
7	7-8	109	16	96	129	225	93	28	41	25	11	9,11,16,17,20,25,26,28,29 (9)	4	4	2	0	0.0
8	8-9	21	88	151	168	319	32	81	10	24	8	none	0	0	0	1(1)	0.3
9	9-10	156	10	152	173	325	66	62	20	25	8	none	0	0	0	0	0.0
10	10-11	48	27	59	66	125	60	9	48	14	11	1,3-9,12 (9)	7	8	6	9,11,19 (3)	2.4
11	11-12	128	7	138	137	275	55	60	20	25	9	2,4,9-16,19 (11)	4	9	3	9 (1)	0.4
12	12-13	113	53	172	242	414	63	74	15	26	6	4 (1)	0	1	0	0	0.0
13	13-1	157	11	155	177	332	85	52	26	31	9	4,14,16 (3)	1	2	1	0	0.0
Total	N/A	1230	N/A	1578	1871	3449	859	N/A	N/A	322	N/A	55	N/A	39	N/A	10	N/A
Average	N/A	95	32	121	144	265	66	50	29	25	10	4	2	3	2	1	0.4
Standard Deviation	N/A	41	25	45	42	83	18	22	16	5	3	4	2	3	2	1	0.7
95% Confidence	N/A	26	15	28	26	51	11	13	10	3	2	3	2	2	1	2	0.9
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18

Columns 1-20 are described from left to right. (1) The animal number. (2-4) Spectral reproducibility was calculated as described in the experimental section. The% match indicates the reproducibility between any two CLC-MS^2 datasets (n and n+1). (5-20) The number of MS^2 spectra remaining and the% of the total number of MS^2 spectra remaining in each data-reduction step (see experimental section). (13) The peptides (listed by the number assigned to the peptide in Table 3-2) and the total number of peptides observed in each sample (shown in parenthesis).

Table 3-2. Peptides and proteolytic processing patterns observed *in vivo*.

Peptide (n)	Animal (n) (no.peptides)	Database-Derived Peptide Sequence	*De Novo* -Derived Peptide Sequence	Position(aa)	Obs (*m/z*)	Mr(expt)	Mr(calc)	Delta	Protein Precursor (no.peptides,% aa sequence coverage, no.animals)
1	1,10 (2)	YGGFM (Met-enkephalin)	[220.1]G[278.1]	100-104,107-111, 136-140,188-192, 212-216,263-267	574.2	573.2	573.2	-0.1	PEA (8,28%,8)
2	11 (1)	YPVEP	PYVKP	118-122	603.9	602.9	603.3	-0.4	
3	10 (1)	YPVEPEEE	YL[226.1]EDPE	118-125	991.4	990.4	990.4	0.0	
4	1,4,10-13 (6)	SPQLEDEAKE (Peptide I$_{1-10}$)	SPKLEDEAKE	198-207	572.9	1143.8	1144.5	-0.8	
5	10 (1)	SPQLEDEAKELQ (Peptide I$_{1-12}$)	[174.0]HNEDEAEGALK	198-209	693.3	1384.7	1385.7	-1,0	
6	4,10 (2)	VGRPEWWMDYQ	FYLDEGWMDYK	219-229	733.9	1465.8	1465.6	0.1	
7	1,2,10 (3)	YGGFL (Leu-enkephalin)	none	232-236	555.6	554.6	555.3	-0.6	
8	1,2,4,10 (4)	YSKEVPEME	FM[168.1]CF[210.2]E	252-260	556.0	1110.1	1110.5	-0.4	
9	5,7,10,11 (4)	RKGPGPGGPGGAGGARGGAGGGPSGD	MGWNVEPYAVKNAGGG[226.1]AD	53-78	687.8	2060.4	2060.0	0.4	Neurogranin (7,33%,7)
10	6,11 (2)	KGPGPGGPGGAGGARGGAGGGP	none	54-75	549.6	1645.8	1644.8	1.0	
11	5-7,11 (4)	KGPGPGGPGGAGGARGGAGGGPSGD	MENGGGAGGRMDRALGGKGN	54-78	635.6	1903.7	1903.9	-0.2	
12	10,11 (2)	GPGPGGPGGAGGARGGAGGGP	[154.1]GGGG[245.0]KGAGGPGGPGH	55-75	759.4	1516.8	1516.7	0.1	
13	11 (1)	GPGPGGPGGAGGARGGAGGGPS	[154.1]GPGAGVKGAGGPGGPGH	55-76	803.0	1604.0	1603.8	0.2	
14	4,11,13 (3)	GPGPGGPGGAGGARGGAGGGPSGD	VEHG[259.1]MAGGPGGPGH	55-78	889.1	1776.2	1775.8	0.4	
15	11 (1)	GPGGPGGAGGARGGAGGGPSGD	[154.1]GGPGGAGKRNAGGG[356.2]	57-78	812.0	1622.0	1621.7	0.3	
16	7,11,13 (3)	ADTGTTDEFIEAGGDIR	none	20-36	870.5	1739.0	1738.8	0.2	Fibrinogen Alpha Chain Precursor (5,5%,4)
17	7 (1)	DTGTTSEFIEAGGDIR	none	21-36	834.4	1666.8	1667.8	-0.9	
18	5 (1)	EFIEAGGDIR	[276.1]LEAGDGLR	27-36	553.4	1104.8	1105.5	-0.7	
19	11 (1)	SPVPDLVPG	S[238.1]VSPVPG	228-236	880.4	879.4	879.5	-0.1	
20	7 (1)	SQLQEGPPEWK	none	240-250	649.8	1297.5	1297.6	-0.1	
21	5 (1)	LVQTQAATTDSDKVDLSIAR	none	13-32	711.0	2130.0	2131.1	-1.1	Fibrinogen Beta Chain Precursor (3,4%,1)
22	5 (1)	TTDSDKVDLSIA (Fibrinopeptide B$_{1-13}$)	[220.1]LMDPGKLS[184.1]	20-31	632.9	1263.8	1263.6	0.1	
23	5 (1)	TDSDKVDLSIAR (Fibrinopeptide B$_{2-14}$)	none	21-32	660.3	1318.7	1318.7	0.0	
24	5 (1)	IAQDNEPEKPVADSETKM	none	526-543	668.2	2001.5	2000.9	0.6	Excitatory Amino Acid Transporter 1 (4,3%,2)
25	5,7 (2)	QDNEPEKPVADSETKM	[171.0]AWGAFHPVVJNPKM	528-543	909.6	1817.1	1816.8	0.3	
26	7 (1)	DNEPEKPVADSETKM	N[244.0][226.1]PVEANSET[259.1]	529-543	845.4	1688.8	1688.8	0.0	
27	5 (1)	EPEKPVADSETKM	none	531-543	730.9	1459.8	1459.7	0.1	
28	6,7 (2)	AKAPAPAAPAAEPQAEAPVAS	none	191-211	957.9	1913.7	1914.0	-0.3	Brain Acidic Membrane Protein (2,13%,2)
29	7 (1)	AKAPAPAAPAAEPQAEAPVASSEQSVAVKE	none	191-220	958.1	2871.2	2871,5	-0.2	
1	2	3	4	5	6	7	8	9	10

Columns 1-10 are described from left to right. (1) The peptide number. (2) The animals (listed by the number assigned to the animal in Table 3-1) and the total number of animals that each peptide was observed in (shown in parenthesis). (3) Database-derived peptide sequences. (4) *De novo*-derived peptide sequences where regions of *de novo*-validated sequences with homology to database-derived sequences are underlined. (5) The amino acid positions of the peptide within the protein precursor. (6) The precursor ion's observed mass-to-charge ratio (*m/z*). (7) The experimental and (8) calculated monoisotopic mass of the peptide. (9) The difference between the experimental and calculated monoisotopic mass. (10) The protein precursors for 29 peptides *in vivo*. The number of peptides observed, the% sequence coverage by amino acid count and the total number of animals each peptide was observed in, are shown in parenthesis.

Peptide Sequencing and Protein Precursor Identification *In Vivo*. Preliminary experiments on microdialysis samples confirmed earlier observations that 3 peptides (i.e.,YGGFM, YGGFL, and SPQLEDEAKE) could be identified by data-dependent MS^2 and database searching when using a precursor ion window of *m/z* 550-600 (animals 1-3). Previous attempts to identify peptides from *in vivo* microdialysis samples using a *m/z* 500-2000 window were unsuccessful[6]. In this work, we found that improving the data-dependent MS^2 detection limit by increasing the injection volume from 1.8 µL to 3.6 µL and decreasing the separation/electrospray flow rate from 20 nL/min to 10 nL/min enabled collection of many MS^2 spectra including those for peptides found in the *m/z* 550-600 range and many novel peptides outside this range. This is most likely due to increases in separation, electrospray and ionization efficiency at lower flow rates.

Twenty-six-3.6 µL injections of CSF microdialysate, collected on-line from animals during basal and depolarization conditions, yielded 121 ± 28 (n = 13) and 144 ± 26 (n = 13) data-dependent MS^2 spectra per chromatogram, respectively, with an *m/z* 550-600 (n = 3) and an *m/z* 500-2000 (n = 10) precursor ion window. (Similar numbers of validated MS^2 spectra were obtained with the two precursor ion windows as shown in Table 3-1; therefore the results are presented together). Figures 3-2 to 3-5 are examples of results obtained *in vivo*.

Results for a single animal are presented in Figures 3-2 and 3-3. Figure 3-2A shows the total ion chromatogram (TIC), Figure 2B shows the MS^2 RIC of the most abundant product ions, and Figure 3-2C shows the RIC of the *m/z* 556 → *m/z* 397 + *m/z* 425 transition for YGGFL. Figure 3-2D shows the MS^2 spectrum for YGGFL. As expected, the a_4 and b_4 ions of YGGFL observed *in vivo* are 5 amu less than the a_4 and b_4 ions of $YGGF_{D5}L$ observed *in vitro* (Figure 1D). Thus, the $YGGF_{D5}L$ standard used to calculate recovery through the microdialysis probe and confirm sensitivity did not interfere with the analysis. Figure 3-3 compares the MS^2 RIC for the most abundant product ions observed *in vivo* for basal (B and D) and depolarization (A and C) conditions for a single animal. Clear differences in the chromatograms suggest detection of a many changes in the chemical environment of the brain extracellular space as a result of this manipulation. In this animal, 96 and 129 MS^2 spectra were collected during basal and depolarization conditions, respectively (Table 3-1). Subtraction of the two datasets (as described in the experimental section) yielded 93 MS^2 spectra that were observed only during depolarization. The other MS^2 spectra were primarily due to contaminants from the microdialysis probe. Database searching revealed that 25 of these MS^2 spectra were significant. From these, 9 MS^2 spectra were validated by identifying the same peptide sequence and protein precursor by Sequest and Mascot with at least two peptides per protein precursor (see experimental section). Thus, 9 peptide sequences and 4 protein precursors were observed during depolarization conditions in this animal (Table 3-2).

Results for all 13 animals are presented in Tables 3-1 and 3-2. Subtractive analysis revealed that 859 of the 3,349 MS^2 spectra collected (25 ± 10% or 66 ± 11 MS^2 spectra per animal) were found only during depolarization (Table 3-1). (Emphasis for data analysis was placed on the spectra observed during depolarization because signaling neuropeptides are expected to be released under these conditions.) Database searching showed that 322 of 3,349 MS^2 spectra (10 ± 2% or 25 ± 3 MS^2 spectra per animal) were significant (i.e., achieved a Sequest ΔCn score greater than 0.1). From these, 17% or 55 of 322 MS^2 spectra (29 were unique) were validated by identifying the same peptide sequence and protein precursor by Sequest and Mascot with at least two peptides per protein precursor (see experimental section). In individual samples, anywhere from zero to 9 of the 29 uniquely validated MS^2 spectra were observed. Six protein precursors were successfully identified with 3 to 28% sequence coverage (Table 3-2). Eighty-three percent or 267 of the 322 significant MS^2 spectra (Sequest) were not validated by a Mascot match, 83% or 267 of 322 were not validated by at least two peptides per protein precursor (i.e., no additional MS^2 spectra were eliminated by requiring at least two peptides per protein precursor) and 17% or 56 of 322 were redundant. Because of redundancy in the animals, 71% or 230 of 322 significant MS^2 spectra corresponded to unique peptide sequences. Twenty-nine of 230 were uniquely validated and 201 of 230 were unique but not validated. Representative MS^2 spectra and sequences for the 29 uniquely validated peptides produced from proteolytic processing of PEA, neurogranin, brain acidic membrane protein, fibrinogen α and β and excitatory amino acid transporter are shown in Figure 3-4 to illustrate the quality of data obtained in this work.

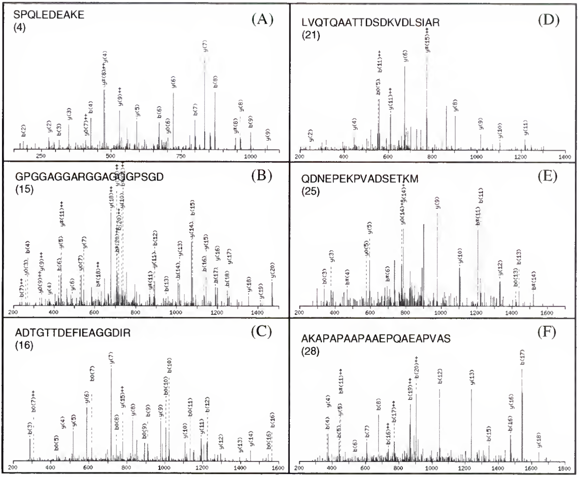


Figure 3-4. Representative MS^2 spectra and sequences for peptides produced from proteolytic processing of PEA (A), neurogranin (B), fibrinogen α (C), fibrinogen β (D), excitatory amino acid transporter (E), and brain acidic membrane protein (F). The number assigned to the peptide in Table 3-1 is listed below the sequence for clarity.

Additional validation was obtained for 62% (18 of 29) of the peptides by showing that the 1st ranked Lutefisk *de novo*–derived partial sequence matched the 1st ranked database-derived sequence better than it matched the 2nd ranked database-derived sequence[60]. In the case of peptide SPQLEDEAKE from PEA observed in animal 10, the MS^2 spectrum yielded a Sequest ΔCn equal to 0.16 and a Mascot score indicating at least homology with 95% probability. The 1st ranked *de novo*–derived partial sequence from Lutefisk was SPKLEDEAKE (Q and K are isobaric in the quadrupole ion trap), while the

2nd ranked sequences from Sequest and Mascot were RRIDEAKE and VNVRSIAGEMGA, respectively. As with the *in vitro* study, this is in excellent agreement with results reported previously where the 1st ranked database-derived sequence was found to be more accurate than the 1st ranked *de novo*-derived sequence[60].

Automated *de novo* sequencing was able to validate only a fraction (62%) of the database-derived sequences (see Table 3-1). For the peptide AKAPAPAAPAAEPQAEAPVAS from brain acidic membrane protein, the MS^2 spectrum yielded a Sequest ΔCn equal to 0.44 and a Mascot score indicating at least homology with 95% probability. No quality sequences were found by Lutefisk, while the 2nd ranked sequences from Sequest and Mascot were EGGEAEAPAAEGGKDEAAGGAA and LTRRIGVGVAVLNRLLY, respectively. In this case, no sequence homology was observed between the database-derived results and the *de novo* sequencing-derived results. The exact reason for this is unknown; however, it is not due to a low S/N MS^2 spectrum. Figure 3-4F shows a high S/N MS^2 spectrum for this peptide. Clearly, automated *de novo* sequencing with Lutefisk is a more stringent criterium for sequence validation than comparing sequences from two different database-searching programs. As with the *in vitro* study, from these results we conclude that validation by automated *de novo* sequencing is confirmatory; however lack of a match between *de novo* sequencing and database searching does not invalidate the sequence identification made by database searching.

The use of a “library” of spectra collected *in vivo* was useful in determining if MS^2 spectra had been incorrectly assigned. Library searching of the 55 MS^2 spectra validated by database searching against the 322 significant MS^2 spectra revealed 10 (3%)

incorrect assignments by the database searching programs as shown in Table 3-1. In the case of SPQLEDEAKE observed in rat 1, the MS^2 spectrum was incorrectly interpreted for the precursor ion at *m/z* 574.4 ($z = +2$) due to a low S/N MS^2 spectrum (particularly in the *m/z* 300-600 region) as shown in Figure 3-5C (despite counts that are ~2-fold higher in C than A). The database-derived sequence was AIKNGWLSEE with a Sequest ΔCn equal to 0.11 and a Mascot score indicating at least homology with 95% probability. The MS^2 spectrum was successfully correlated with SPQLEDEAKE by searching the *in vivo* library for matching MS^2 spectra as shown by the difference spectrum in Figure 3-5B. Library searching was performed using Xcalibur (ver. 1.2) in conjunction with NIST MS Search (ver. 1.7). A simple similarity search, where the algorithm weights the ions in the MS^2 spectra by mass to find MS^2 spectra that are similar to the query MS^2 spectrum, was selected with the default NIST MS Search options. The MS^2 spectrum for SPQLEDEAKE from rat 1 (5C) matched the significant MS^2 spectra from the other rats in the *in vivo* library with less than a 38% difference in the product ion intensities of the MS^2 spectra. For example, the MS^2 spectrum for rat 12, with a Sequest ΔCn equal to 0.23, is shown in 5A and the difference spectrum (i.e., 5A - 5C) is shown in 5B. Based on the frequency of our observations (i.e., SPQLEDEAKE was observed in 6 of 13 animals while AIKNGWLSEE was observed only once), the Sequest ΔCn scores and the MS^2 spectrum for the synthetic peptide SPQLEDEAKE[6], the correct assignment for the MS^2 spectrum in (5C) is SPQLEDEAKE rather than AIKNGWLSEE. Co-release of other PEA-derived peptide sequences provides additional evidence in support of this argument. We found that library searching 'rescued' a few incorrectly assigned MS^2

spectra and thus slightly improved the reproducibility of peptide sequencing and protein precursor identification *in vivo.*

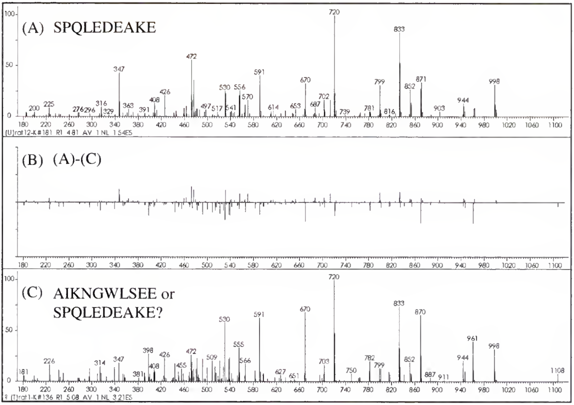


Figure 3-5. MS^2 spectra (A and C) from the *in vivo* library and the difference MS^2 spectrum (B) for SPQLEDEAKE observed in two rats. Database-searching programs incorrectly assigned the sequence AIKNGWLSEE to the MS^2 spectrum in (C).

A summary of the *in vivo* results is shown in Figure 3-6 to illustrate the data-reduction strategy described in the experimental section. A decrease in the number of MS^2 spectra remaining at each data-reduction step clearly indicates successively more stringent criteria for sequence validation. Of the MS^2 spectra collected *in vivo*, there were 3,449 MS^2 spectra remaining after data-reduction step 1, 859 depolarization-specific MS^2 spectra (29 ± 10%) remaining in step 2, 322 significant MS^2 spectra (10 ± 2%) remaining in step 3, 55 database-validated MS^2 spectra (2 ± 1%) remaining from step 4 (29 were unique), 39 *de novo*-validated MS^2 spectra (2 ± 1%) remaining in step 5 and 10 incorrectly assigned MS^2 spectra (0.4 ± 0.9%) remaining in step 6. Subtraction of the 55 validated MS^2 spectra and 10 incorrectly assigned MS^2 spectra from the 322 significant MS^2 spectra yielded 257 MS^2 spectra that did not meet the criteria for validation. From the remaining 257 MS^2 spectra, 56 were redundant and 201 were unique (i.e., they did not match other MS^2 spectra in the library). In order to prevent false-positives, further validation will be required to confirm the corresponding peptide sequences and protein precursors.

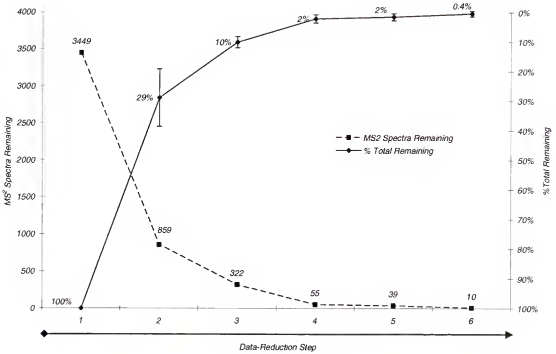


Figure 3-6. A summary of the *in vivo* results illustrating the six step data-reduction strategy described in the experimental section. The number of MS^2 spectra remaining and the% of the total number of MS^2 spectra remaining are plotted vs. the data-reduction step (see experimental section).

No MS^2 spectra collected during basal conditions could be validated in any of the animals. This suggests that the peptide sequences we observed during depolarization are not simply the products of neuronal injury caused by the microdialysis probe as these would have been observed during basal conditions (i.e., closer to the time that the dialysis probe was placed into the brain). This also suggests that with anesthetized animals and the sampling methods used, not enough peptide can be recovered during basal conditions to routinely measure peptides and proteins present in the extracellular space. However, the use of K^+-induced depolarization can raise levels due to secretion and other processes to allow peptide sequencing and protein precursor identification *in vivo*. Significantly, subtraction of poor quality MS^2 spectra (i.e., basal conditions) from high quality MS^2 spectra (i.e., depolarization conditions) prior to database searching did not result in a loss of information. Rather, subtractive analysis simply reduced the number of MS^2 spectra requiring database searching from 3,349 to 859 and significantly reduced processing time ~ 4-fold. This was confirmed by obtaining the same validated MS^2 spectra without subtractive analysis prior to database searching for several animals (n = 3).

Spectral reproducibility for *in vivo* samples was determined by comparing data sets from different animals using subtractive analysis. This evaluation revealed that on average 32 ± 15% or 132 ± 22 MS^2 spectra collected per animal (specific to depolarization) were the same in comparing samples from any two animals. Spectral reproducibility *in vivo* was significantly lower than the 75% found for the tryptic digests discussed above. Interanimal variability can result not only from real differences between animals, but also experimental differences such as probe placement. We estimate the peptide concentrations to be 100-2000 pM[65] corresponding to 360-7200 attomoles injected on-column *in vivo*. Because the estimated peptides concentrations were 1 nM corresponding to 400 attomoles injected on-column *in vitro*, it is improbable that the increased variability *in vivo* was simply due to lower levels of peptides in the dialysate than in the tryptic digest.

Comparison of Peptide Sequencing and Protein Precursor Identification *In Vitro* and *In Vivo*. The results indicate that proteomics tools can be applied, with some degradation in performance, to dialysate samples despite the difficulties of limited samples and unknown protease specificity. A comparison of the results from the *in vitro* experiments to the *in vivo* experiments provides a sense of the limitations of the *in vivo* approach Spectral reproducibility was 75 ± 18% *in vitro* and 32 ± 15% *in vivo*. This reduction in reproducibility was most likely due to inter-animal variability from real differences between animals and from experimental differences such as probe placement. Reproducibility of peptide sequencing and protein precursor identification was also markedly reduced in the *in vivo* experiments. For the *in vitro* experiments, 7 of 8 peptides were sequenced in 3 samples. In contrast, 0 to 9 peptides were sequenced in 13 animals. It is likely that greater reproducibility can be obtained *in vivo* by optimizing the gradient steepness parameter[65], precursor ion windows and other data-dependent MS^2 scan functions (e.g. collision energy) for protein sequence coverage and by minimizing probe placement variability and, if possible, inter-animal variability.

In vivo samples were analyzed using both a *m/z* 500-2000 precursor ion window and a *m/z* 550-600 window. Using a 30-fold wider scan range, one would expect the presence of many more peptides to detect. With the narrower scan range, it was expected that greater sensitivity would be achieved allowing more low level peptides to be detected within this range. Since similar numbers of MS^2 spectra were collected (where the number of MS^2 spectra ~ analysis time / 'triple play' cycle time) and validated with both approaches, it appears that the two effects compensate each other. Thus, it is expected that maximal coverage of the dialysate proteome would be achieved by multiple injections of samples with narrow scan ranges (e.g., *m/z* 50). However, this is a time-consuming process. The wider scan range allows peptides across the entire *m/z* range to be detected; albeit only the higher level peptides. Presumably the use of the wider scan function detects higher level peptides which may be of greater interest as possible functional peptides.

Single Peptide Sequence Validation and Protein Precursor Identification. Single peptide sequences may be used for protein identification provided they are sufficiently validated. While single peptide sequences are often used for protein identification, they are seldom validated despite recent acknowledgement of false-positives using widely accepted scoring criteria[57,58]. Moreover, no rigorous examination of the false-positive rate for protein identification as a function of the number of peptides sequenced has been performed. The six step data-reduction strategy (Figure 3-6) described in the experimental section yielded only 2% or 55 of 3,349 MS^2 spectra (29 were unique) that were validated by database searching. Closer inspection of data-reduction steps 3 to 5 are shown in Figure 3-7 to illustrate the interaction of multiple selection criteria as described in the experimental section. In step 4, MS^2 spectra remaining were considered validated by database searching if the same peptide sequence and protein precursor were identified by Sequest and Mascot (step 4a), and if at least two peptides per protein precursor were found (step 4b). The diagram in Figure 3-7 shows the overlap between MS^2 spectra remaining at different points of data-reduction steps 3 to 5. The overlap between the 1st ranked Sequest-derived sequences (322 MS^2 spectra received a ΔCn score greater than 0.1 in step 3) and the 1st ranked Mascot-derived sequences (19% or 62 of 322 MS^2 spectra indicated homology with greater than 95% probability) was 17% or 55 of 322 MS^2 spectra (step 4a). All of these overlapping MS^2 spectra, 17% or 55 of 322, were found to correspond to at least two peptides per protein precursor (step 4b). Thus, Mascot effectively reduced false-positives, but no additional MS^2 spectra were eliminated by requiring at least two peptides per protein precursor. This suggests that the selection criteria in data-reduction step 4a were more stringent than the selection criteria in step 4b. In other words, it was unnecessary to require at least two peptides per protein precursor if the 1st ranked Sequest- and Mascot-derived sequences matched each other. Automated *de novo* sequencing provided further validation for 12% or 39 of 322 MS^2 spectra (step 5); however, in many cases it failed to validate MS^2 spectra despite validation by database searching programs (step 4). Thus, the selection criterium in step 5 was more stringent than the selection criteria in step 4. Library searching (step 6) revealed that only 3% or 10 of 322 significant MS^2 spectra (those that received a Sequest a ΔCn score greater than 0.1 in step 3) were incorrectly assigned by database searching. Again, the selection criterium in step 6 was more stringent than the selection criteria in step 3. While it is possible that many of the 201 unique but not validated MS^2 spectra were properly assigned by Sequest, greater confidence in peptide sequencing and protein precursor identification at the attomole level was obtained by applying successively more stringent data-reduction steps. Therefore, comparing sequences obtained from two different database-searching programs (the 2nd more stringent than the 1st) and investigating additional validation by one *de novo* sequencing program was our preferred method of single peptide sequence validation and protein precursor identification. This is preferable to sequence validation by peptide synthesis or manual sequence interpretation, which are expensive and time-consuming, respectively.

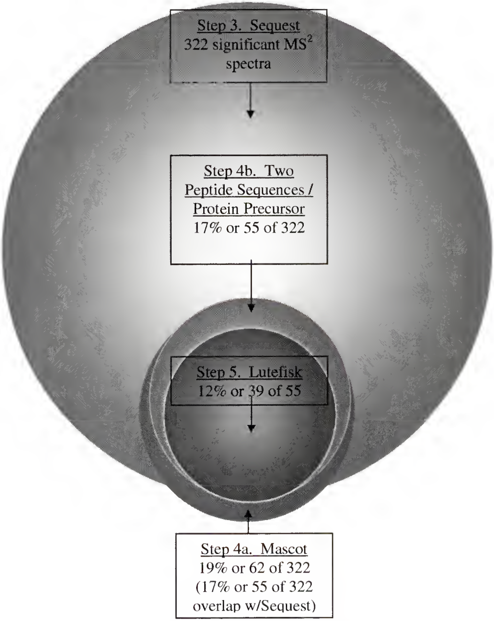

Step 3. Sequest
322 significant MS2 spectra
Step 4b. Two Peptide Sequences / Protein Precursor
17% or 55 of 322
Step 5. Lutefisk
12% or 39 of 55
Step 4a. Mascot
19% or 62 of 322
(17% or 55 of 322 overlap w/Sequest)

Figure 3-7. Diagram showing the interaction of multiple selection criteria and the overlap between MS^2 spectra remaining at data-reduction steps 4 and 5 (see experimental section).

Conclusion

***In Vivo* Microdialysis-CLC-MS^2.** We have shown that *in vivo* microdialysis-CLC-MS^2 enables peptide sequences and protein precursors to be correlated to K^+-induced release without *a priori* knowledge of the molecules involved. This was accomplished using 3.6 µL injections of CSF microdialysate, collected on-line, containing attomole amounts of peptides at picomolar concentrations produced from unknown proteolytic processing events. We have overcome the challenges arising from mass- and concentration limited dialysate samples by the combination of miniaturized CLC columns with integrated electrospray emitters, optimized gradients, tandem mass spectrometry, and proteomics methods.

A rigorous data-reduction strategy was applied to minimize the possibility of false-positive peptide sequences and protein precursor identifications. Further development and integration of these technologies is expected to enable correlation of other previously-unknown molecules to specific behavioral or physiological states of a living animal at the zeptomole level. This provides a powerful approach to asking and answering many important biological questions.

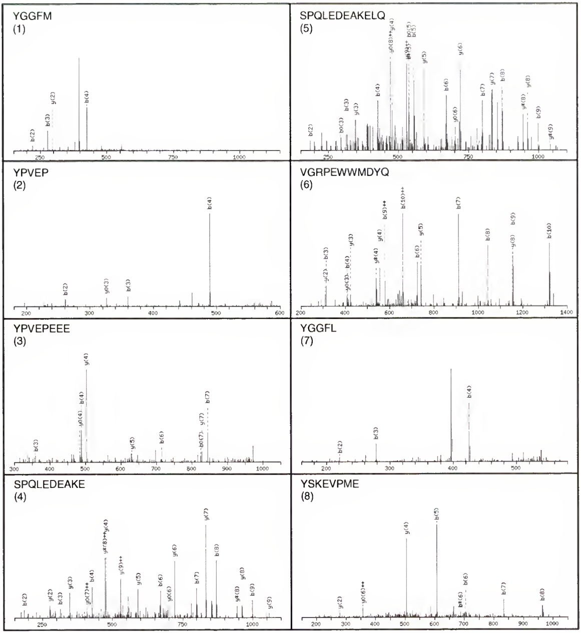
CSF Proteomics. The CSF proteome is relatively unexplored because CSF samples are both mass- and concentration-limited. We estimate that the peptides observed in this work were between 100-2000 pM in dialysate. Thus, traditional CLC-MS^2 employing *in vitro* proteolysis at femtomole levels could not be applied. There are approximately N/10 tryptic peptides compared with N(N+1)/2 possible subsequences for a protein of N amino acids[66]. This causes a nightmarish 100-1000 fold increase in processing time with present database searching programs. For example, PEA (269 aa) contains only 25 tryptic peptides (21 of *m/z* 500-2000) but 21,528 possible subsequences! Post-translational modifications (PTMs) common to neuropeptides (e.g., N-terminal pyroglutamate) further complicate sequence analysis with an exponential increase in processing time. Clearly, sequencing peptides from unknown proteins and discovering proteolytic processing patterns are more challenging problems than identifying unknown proteins from their tryptic peptides. None of the validated MS^2 spectra observed in this work corresponded to sequences with PTMs. This is likely because most PTMs are non-quantitative, suggesting that peptides containing PTMs are at the low attomole level in CSF and below the current detection limit of data-dependent MS^2 scans. However, further investigation of selected PTMs could be investigated using time-segmented MS^2 scans with low attomole detection limits[6].

The information contained in proteolytic processing is crucial for understanding biological function. Over 250,000 papers on neuroactive peptides have been published since 1965[67], and the paper describing the discovery of Met- and Leu-enkaphalin has been cited over 2840 times (http://isiknowledge.com)! According to the MEROPS protease database (http://merops.sanger.ac.uk/index.htm), 241 proteases have been observed in *Rattus norvegicus* and 539 in *Homo sapiens*[68]. The physiological roles of the peptides discovered in this work and the proteases involved in their production remain to be determined. However, co-release of 2 known neuropeptides, the opioids Met-enkephalin and Leu-enkephalin, suggests that important inter-neuron signaling mechanisms and protein degradation pathways[69] have been discovered. In other words, the novel peptides presented herein are likely to be bioactive and/or involved in the regulation of Met- and

Leu-enkephalin in the ECF. Significantly, all of the peptide and protein sequences reported in this work are also found in the human proteome.

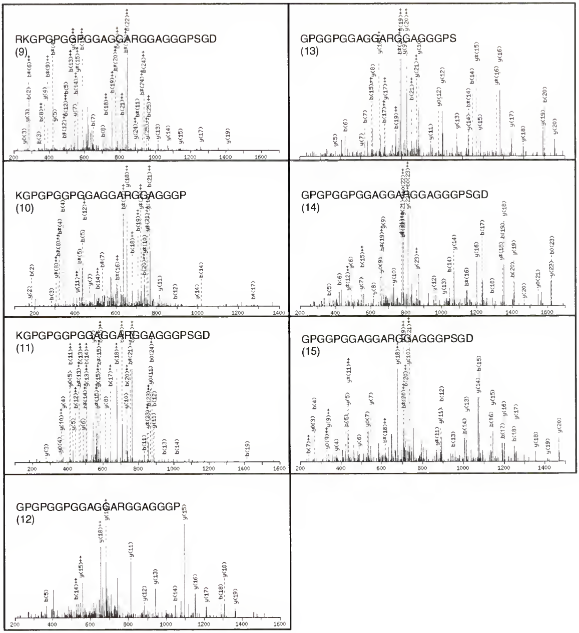
Supporting Information

Figure 3-8. MS^2 spectra and sequences for peptides produced from proteolytic processing of PEA (A), neurogranin (B), fibrinogen α (C), fibrinogen β (D), excitatory amino acid transporter (E) and brain acidic membrane protein (F). The number assigned to the peptide in Table 3-2 is listed below the sequence for clarity.



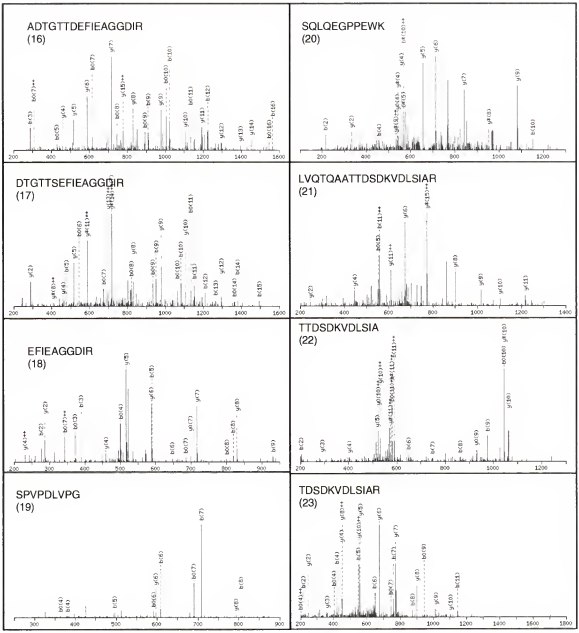
(A)

Figure 3-8. Continued.



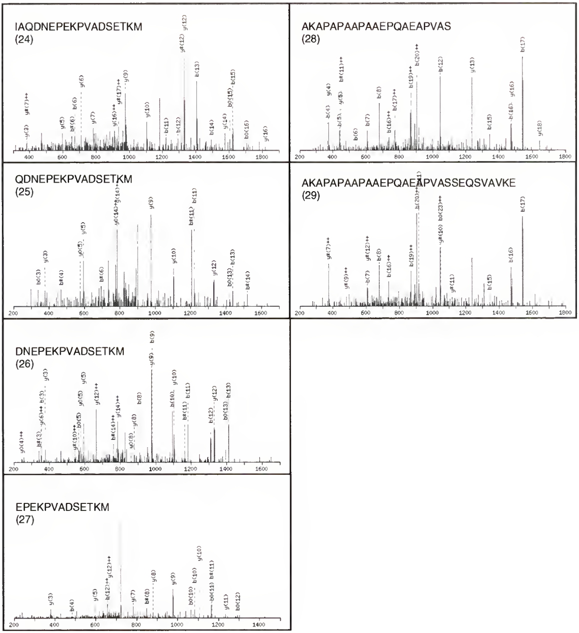
(B)

Figure 3-8. Continued.



(C and D)

Figure 3-8. Continued.



(E and F)

Figure 3-8. Continued.

Table 3-3. Summary of *in vitro* CLC-MS^2 results. Columns 1-13 are described from left to right. (1) The sample number. (2-4) Spectral reproducibility was calculated as described in the experimental section. The% match indicates the reproducibility between any two CLC-MS^2 datasets (n and n+1). (5-13) The number of MS^2 spectra remaining and the% of the total number of MS^2 spectra remaining in each data-reduction step (see experimental section). (8) The peptides (listed by the number assigned to the peptide in Table 3-4) and the total number of peptides observed in each sample (shown in parenthesis).

Sample (n)	***Spectral Reproducibility***			**Data-Reduction Step**								
				1	**3**		**4**		***5***		**6**	
	Samples	*n-(n+1)*	*% Match*	*Total*	*No.*	*% Total*	*No.*	*% Total*	*No.*	*% Total*	*No.*	*% Total*
1	1-2	39	68	122	23	19	1-3,5-8,YGGF$_{D5}$L (8)	21	6	6	0	0
2	2-3	24	82	134	18	13	1-3,5-8,YGGF$_{D5}$L (8)	33	6	5	0	0
3	3-1	36	74	139	21	15	1-8,YGGF$_{D5}$L (9)	25	7	6	0	0
Total	N/A	99	N/A	395	62	N/A	25	N/A	19	N/A	0	N/A
Average	N/A	33	75	132	21	16	8	26	6	6	0	0
Standard Deviation	N/A	8	7	9	3	3	1	7	1	0	0	0
95% Confidence	N/A	20	18	22	6	7	1	16	1	1	0	0
1	2	3	4	5	6	7	8	9	10	11	12	13

Table 3-4. Peptides and proteolytic processing patterns observed *in vitro*. Columns 1-10 are described from left to right. (1) The peptide number. (2) The sample (listed by the number assigned to the sample in Table 3-3) and the total number of samples that each peptide was observed in (shown in parenthesis). (3) Database-derived peptide sequences. (4) *De novo*-derived peptide sequences where regions of *de novo*-validated sequences with homology to database-derived sequences are underlined. (5) The amino acid positions of the peptide within the protein precursor. (6) The precursor ion's observed mass-to-charge ratio (*m/z*). (7) The experimental and (8) calculated monoisotopic mass of the peptide. (9) The difference between the experimental and calculated monoisotopic mass. (10) The protein precursors for 8 peptides *in vitro*. The number of peptides observed, the% sequence coverage by amino acid count and the total number of samples each peptide was observed in, are shown in parenthesis.

Peptide (n)	Sample (n) (no.peptides)	Database-Derived Peptide Sequence	*De Novo*-Derived Peptide Sequence	Position(aa)	Obs (*m/z*)	Mr(expt)	Mr(calc)	Delta	Protein Precursor (no.peptides, % aa sequence coverage, no.samples)
1	1-3 (3)	FESNFNTQATNR	[276.1]LAKTNFAMTR	34-45	715.3	1428.7	1427.6	1.0	Lysozyme (3,29%,3)
2	1-3 (3)	NTDGSTDYGILQINSR	[262.1]LLKLLGYDTVNSR	46-61	878.0	1754.0	1752.8	1.2	
3	1-3 (3)	GTDVQAWIR	none	117-125	523.8	1045.7	1044.5	1.1	
4	3 (1)	DFPIANGER	[262.1]PLANGER	18-26	509.8	1017.5	1017.5	0.0	Carbonic Anhydrase II (3,11%,3)
5	1-3 (3)	ALVYGEATSR	none	47-56	534.4	1066.7	1065.6	1.2	
6	1-3 (3)	VGDANPALQK	VGDAN[168.1]LKK	148-157	506.8	1011.7	1011.5	0.1	
7	1-3 (3)	DGKGDVAFVK	DGAGGD[154.1]YVK	200-209	518.4	1034.7	1034.5	0.2	Conalbumin (2,3%,3)
8	1-3 (3)	AQSDFGVDTK	[199.1]SDFRD[229.0]	270-279	534.5	1067.0	1066.5	0.5	
1	2	3	4	5	6	7	8	9	10

CHAPTER 4
STUDYING ENDOGENOUS NEUROPEPTIDE PROCESSING BY *IN VIVO* MICRODIALYSIS-CLC-MS^2

Introduction

Proteolytic processing is a complex type of post-translational modification (PTM) that generates biologically active peptides with important cell signaling properties. Tissue-specific processing[45] of large precursor proteins is performed by proteases, endo- and exopeptidases, during intracellular transport from the rough endoplasmic reticulum where the proteins are folded by molecular chaperones to the Golgi apparatus where various PTMs are performed[70]. Further proteolytic processing occurs during stimulation and exocytotic release from a heterogenous population of secretory vesicles (SVs) into the ECF. Understanding the cascades of limited proteolytic steps by several highly specific proteases as a mechanism to increase biological diversity is important in cell signaling and in the production of recombinant proteins[71].

Differential processing by various proteases produces a complex mixture of peptide intermediates and bioactive peptides. Proteases typically cleave at mono- and di-basic amino acid residues (e.g., R, K, KK and RK) flanking the N- and C-terminus of the peptide sequence within the protein precursor. The general mechanism for proteolytic processing is: 1) endopeptidase cleavage at the C-terminal side of mono- and di-basic amino acid residues, 2) carboxypeptidase cleavage of the newly exposed C-terminal mono- and di-basic amino acid residues, 3) peptidyl-amidating monooxygenase (PAM) amidation of C-terminal glycine residues and 4) aminopeptidase cleavage of one or more N-terminal residues to yield the shortened mature peptides[72].

Prediction of protease cleavage sites and proteolytic processing products[73] is precluded by a number of unknowns. This includes the consensus sequences and subcellular locations of the proteases involved, the kinetics of proteolysis and the tertiary and quaternary structures of the proteases and substrate. Therefore, experimental approaches are required to identify peptide products and intermediates in the various proteolytic processing pathways.

Proteolytic processing of preproenekphalin A (PEA) has been extensively studied in brain tissue[45,74-78] and purified chromaffin granules of bovine adrenal medullas (BAM)[79-94]. These studies revealed that PEA-processing occurs via successive proteolytic events by various proteases found in the brain such as preprotein convertase 1/3 and 2 (PC1/3 and PC2) to generate biologically active peptides such as the neuropeptides Met- and Leu-enkephalin. However, a problem with these studies is that no discrimination was made between intra- and extracellular-peptides. More recently, *in vivo* microdialysis has enabled PEA-derived enkephalins to be studied in the extracellular fluid (ECF)[21,47,95-98]. However the complete protelytic processing pathway for PEA remains unknown.

The objective of this work is to describe the biological significance of peptide sequences and protein precursors discovered in the ECF by *in vivo* microdialysis-capillary liquid chromatography (CLC)-tandem mass spectrometry (MS^2). Previously, we demonstrated that CLC columns with integrated electrospray emitters operated at low nanoliter/min flow rates interfaced to quadrupole ion trap mass spectrometers could be used to monitor and discover endogenous peptides at attomole levels in microdialysis samples[65]. In chapter 3, proteolytic processing (i.e., the production of peptides by the action of proteases on proteins) was explored by *in vivo* microdialysis-CLC-data-dependent-MS^2. Here, we discuss proteolytic processing, as a function of the putative proteases involved in the cleavage of each peptide sequence from its protein precursor, of PEA and 5 other proteins whose peptide products were measured simultaneously. We also present preliminary evidence for biological activity of a novel PEA-derived peptide discovered by this approach These studies are expected to be particularly important for microdialysis-based proteomics applications in neuroscience.

Experimental

The experimental section was described previously in chapter 3.

Results and Discussion

Overview. A brief description of the results obtained *in vivo* is given here to facilitate a discussion of the biological significance of the peptide sequences and protein precursors discovered in this work. The development, evaluation and application of a proteomics-based data-reduction strategy in conjunction with the *in vivo* microdialysis-CLC-MS^2 system for discovering peptides were described previously (Chapter3).

***In Vivo* Microdialysis-CLC-MS^2.** A total of 3,349 MS^2 spectra were collected from 13 different animals under basal conditions (with artificial CSF perfusing the probe) and during localized depolarization evoked by infusion of a high K^+, low Na^+ solution through the dialysis probe. Subtractive analysis revealed a total of 322 MS^2 spectra that were observed only during depolarization and received a significant score by database searching. From these MS^2 spectra, 29 peptide sequences and 6 protein precursors were identified using the database searching programs Sequest and Mascot. Figure 4-1 is an example of results obtained *in vivo* during depolarization-induced release of peptides into the ECF.

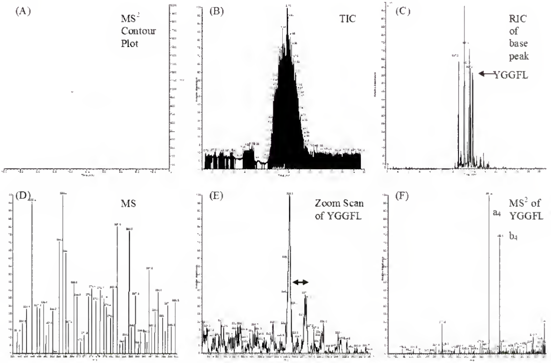


Figure 4-1. *In vivo* microdialysis-CLC-MS^2 analysis of CSF collected from the rat striatum during depolarization-induced release of peptides into the ECF. Figures 4-1A-C illustrate the separation of peptides and product ions during CLC-MS^2 analysis of dialysate. Figure 4-1A shows a MS^2 contour plot of the most abundant product ions observed during a data-dependent MS^2 experiment, Figure 1B shows the total ion chromatogram (TIC) and Figure 1C shows the base peak reconstructed ion chromatogram (RIC) for the most abundant product ions. An arrow indicates the peak corresponding to YGGFL. Figures 4-2D-F illustrate how data-dependent MS^2 spectra were collected in the 'triple play' scan mode (MS, zoom, MS^2). Figure 4-2D shows a MS scan where precursor ion for YGGFL at *m/z* 556 exceeded the threshold for data-dependent MS^2 analysis. Figure 4-2E shows a zoom scan for charge state determination of the *m/z* 556 precursor ion in (D). A $\Delta m/z = 1.0$ between the ^{12}C and ^{13}C isotopes yields $z = +1$ and a $[M+H]^+$ precursor ion. Figure 4-2F shows the data-dependent MS^2 spectrum of the singly charged precursor ion at *m/z* 556 in (D) with the characteristic a_4 and b_4 product ions (*m/z* 397 and *m/z* 425, respectively) of YGGFL.

Depolarization-Specific Proteolytic Processing *In Vivo*. Stimulation by infusion of a high K^+, low Na^+ solution causes neuronal depolarization, Ca^{2+} influx, fusion of SVs with the plasma membrane, and release of the contents of SVs into the ECF by exocytosis[99]. Therefore, our subtractive analysis method was expected to reveal only neuropeptides that were stored in neurons and released under de-polarizing conditions. Co-release of the known neuropeptides, the opioids Met- and Leu-enkephalin, along with 6 other PEA-derived peptides were found by this approach. Depolarization also caused increases in other peptides including neurogranin-, fibrinogen α and β- and excitatory amino acid transporter 1-derived peptides.

The presence of these other peptides during K^+-induced depolarization is of interest as not all of these were expected to be released from synaptic vesicles. One possibility is that depolarization raised the levels of proteases in the extracellular space which degraded the extracellular domains of synaptic or surface proteins and raised the levels of peptides in the ECF. Another possibility is that depolarization caused cell injury resulting in disruption of the blood brain barrier (BBB) and subsequent elevation of the levels of proteases, proteins and peptides in the ECF. Irrespective of the mechanism that raised the levels of peptides in the ECF during K^+-induced depolarization, proteolytic processing can be described as a function of the putative proteases involved in the cleavage of each peptide sequence from its protein precursor.

The putative proteases involved in cleavage of 29 peptide sequences (27 novel) from 6 protein precursors are shown in Table 4-1. Proteases were identified for the N- and C-terminus cleavage sites of each peptide by searching the MEROPS protease database (http://merops.sanger.ac.uk/index.htm) which contains greater than 241 proteases for *Rattus norvegicus* and 539 for *Homo sapiens*[100]. Serine-, cysteine-, thiol-, aspartic-, metallo-, amino- and carboxy-endopeptidases are all represented in Table 4-1. Many of the proteases in Table 4-1 have not yet been found in the rat brain; however, these are listed in Table 4-1 because genes encoding proteases are high conserved[43] and many novel proteases with homology to these known proteases are expected to be discovered in forthcoming years. Proteases in bold are specific to 3 or more amino acids in the peptide sequence (starting one amino acid past the cleavage site), while other proteases are specific only to the 2 amino acids spanning the cleavage site. While we cannot unambiguously determine which proteases were involved in cleavage of each peptide and protein by our approach, inspection of Table 4-1 by protease specificity, where bold proteases are the most significant, provides clues towards this end.

Table 4-1. Putative proteases involved in the proteolytic processing patterns observed *in vivo*. Proteolytic processing patterns are described as a function of the putative proteases involved in N- and C-terminus bond cleavage of each peptide sequence from its protein precursor. Proteases in bold are specific to 3 or more amino acids in the peptide sequence (starting one amino acid past the cleavage site).

Peptide (n)	Position(aa)	Peptide Sequence	Protein Precursor (no.peptides, % aa sequence coverage, no.animals)	Putative Peptidases from MEROPS Database			
				N-terminus Cleavage Point	N-Terminus Peptidases	C-terminus Cleavage Point	C-Terminus Peptidases
1	100-104,107-111,136-140,188-192,212-216 263-267	YGGFM (Met-enkephalin)	Preproenkephalin A (8 28% 8)	R/Y	**Proprotein convertase 1,2**	M/R	Cathepsin X
				K/Y	Polyporopepsin Neurosporapepsin, Hyicolysin, Tryptase alpha, u-Plasminogen activator; Angiotensin converting enzyme multipeptidase	M/K	**Carboxypeptidase M** Pappalysin-1, Kallikrein 1
2	118-122	YPVEP		R/Y	**Proprotein convertase 1,2**	P/E	Unknown
3	118-125	YPVEPEEE				E/A	Gastricsin, Cathepsin E Rhizopuspepsin Phytepsin Nepenthesin Soytalidopepsin B Aspergillopepsin II Staphopain A Aerugonolysin Insulysin, Penicillolycin, Deuterolysin Glutamyl aminopeptidase, Glycyl aminopeptidase Proteinase K Mername-AA054 peptidase Archaean proteasome beta, Prohead proteinase
4	198-207	SPQLEDEAKE (Peptide I_{1-10})		R/S	**Proprotein convertase 1,2** 4 5 and 7, Yapsin 1, Separase Conn Kallikrein 1,mK1,mK6 mk9 rK1 rK10, Snake venom factor V activator; Complement factor 1,factorB,component2, Thrombin, Coagulation Factor C, Protein C, Falvivirin, PACE 4 proprotein convertase tripeptidyl-peptidase C	E/L	Matrilysin
5	198-209	SPQLEDEAKELQ (Peptide I_{1-12})				Q/K	Canditropsin, Bontoxilysin, CPG70 carboxypeptidase
6	219-229	VGRPEWWMDYQ		R/V	Cycstinyl aminopeptidase, dipeptidyl peptidase III, Easter endopeptidase, Coagulation factor VIIa XIa,XIIa Plasma kallikrein Acrosin, Hepatocyte growth factor activator, u- and t Plasminogen activator **Intestinal Arg-specific endopeptidase**	Q/K	Canditropsin Bontoxilysin, **CPG70 carboxypeptidase**
7	232-236	YGGFL (Leu-enkephalin)		R/Y	**Proprotein Convertase 1,2**	L/K	Pepsin A Renin, Cathepsin D, Rhizopuspepsin Barrier Pepsin, Felin immunodeficiency virus retropepsin, Carboxypeptidase M, CPG70 Carboxypeptidase
8	252-260	YSKEVPEME		S/Y	Halylysin, Lysosomal dipeptidase I **Acylaminoacyl-peptidase**	E/K	Oligopeptidase A, CPG70 Carboxypeptidase Peptido-Lys-metalloendopeptidase

Table 4-1. Continued.

9	53-78	RKGPGPGGPGGAGGARGGAGGGPSGD	Neurogranin (7,33%,7)	G/R	Cathepsin B X, Carboxypeptidase B,U, foot-and-mouth disease virus L proteinase, Aut2 peptidase Leucolysin	None	N/A
10	54-75	KGPGPGGPGGAGGARGGAGGGP		R/K	Proprotein convertase 2 Separase, Carboxypeptidase U Pitrilysin, Nardilysin, Hyicolysin, **Tryptase alpha**, Granzyme K, Complement factor D, Complement component activated C1s, Mannan-binding lectin-associated serine protease 1,2. Protein C, Intestinal Arg-specific endopeptidase	P/S	Ig A1-specific serine endopeptidase
11	54-78	KGPGPGGPGGAGGARGGAGGGPSGD				None	N/A
12	55-75	GPGPGGPGGAGGARGGAGGGP		K/G	Proprotein converatase 2, Foot-and-mouth disease virus L proteinase	P/S	Ig A1-specific serine endopeptidase
13	55-76	GPGPGGPGGAGGARGGAGGGPS				S/G	**Nuclear inclusion a-endopeptidase**
14	55-78	GPGPGGPGGAGGARGGAGGGPSGD				None	N/A
15	57-78	GPGGPGGAGGARGGAGGGPSGD		P/G	Helylysin, **Cytophagelysin** Membrane Pro-X carboxypeptidase, Prolyl oligopeptidase **Dipeptidyl peptidase II**		
16	20-36	ADTGTTDEFIEAGGDIR	Fibrinogen Alpha Chain Precursor (5,5% 4)	A/A	Saccharolysin, Peptidyl dipeptidase Dcp, Gametolysin, Peptidase T Bontoxilysin, Glycyl aminopeptidase, Beta-Ala-His dipeptidase, Lysosomal dipeptidase I, Membrane Pro-X-carboxypeptidase, Pancreatic Elastase Cytomegalovirus assemblin, Signal peptidase I, Dipeptidyl-peptidase II, C-terminal processing protease-1	R/R	Proprotein convertase 2 Omptin Separase, Thimet oligopeptidase, Carboxypeptidase U Insulysin Nardilysin, Penicillolysin, Dipeptidyl-peptidase III, BlaR1-peptidase, **Kallikrein rK1** mK1, mK6 mK9 Intestinal Arg-specific endopeptidase
17	21-36	DTGTTSEFIEAGGDIR		A/D	Procollagen C-proteinase Flavastacin		
18	27-36	EFIEAGGDIR		S/E	Cathepsin X		
19	228-236	SPVPDLVPG		M/S	Papain, Serralysin Met-X dipeptidase Lysosomal dipeptidase I	G/S	Cathepsin X, Leishmanolysin, Aeruginolysin, Halyllysin, Ecarin, Deuterolysin, Kallidin-releasing enzyme, *Physarum* serine carboxyl protease, Archaean proteasome-beta component Meprin A peptidase complex
20	240-250	SQLQEGPPEWK		K/S	Trypsin I	K/A	Proprotein convertase 1,2, Penicillopepsin Rhizopuspepsin, Rhodotorulapepsin Bontoxilysin, Plasma Kallikrein
21	13-32	LVQTQAATTDSDKVDLSIAR	Fibrinogen Beta Chain Precursor (3.4%,1)	S/L	Lysosomal dipeptidase I, Brevilysin L6 Hornlysin, Acutolysin Fragilysin, Rous sarcoma virus retropepsin	R/G	Proprotein converatse 1,2, **Thrombin** Protein C Clotting Enzyme, Flavinn, Furin, Peptidyl-glycinamdiase Geunidino benzoatase, **Venombin AB** Angiotensin converting enzyme multipeptidase, Ancrod Crotalase Batroxobin Trypsin 1, PoFibs endopeptidase Deuterolysin, Hemorrhagic metalloproteinase e Kistomin Bilitoxin, Aeruginolysin Gingipain R, Clostripain, Poliovirus-type picornain 2a Grapevine-fanleaf-type nepovirus pecornain 3c
22	20-31	TTDSDKVDLSIA (Fibrinopeptide B_{1-13})		A/T	Presenelin 1, Tymovirus endopeptidase	A/R	Carboxypeptidase B,M N,T E,U, CPG70 carboxypeptidase, **Dipeptidyl-peptidase 7** Carboxypeptidase D multipeptidase
23	21-32	TDSDKVDLSIAR (Fibrinopeptide B_{2-14})		T/T	Unknown	R/G	Proprotein converatse 1 2 **Thrombin** Protein C Clotting Enzyme, Flavinn Furin, Peptidyl-glycinamdiase, Geunidino benzoatase, **Venombin AB** Angiotensin converting enzyme multipeptidase Ancrod, Crotalase Batroxobin, Trypsin 1 PoFibs endopeptidase Deuterolysin Hemorrhagic metalloproteinase e, Kistomin, Bilitoxin, Aeruginolysin, Gingipain R Clostripain Poliovirus-type picornain 2a, Grapevine-fanleaf-type nepovirus pecornain 3c

Table 4-1. Continued.

24	526-543	IAQDNEPEKPVADSETKM	Excitatory Amino Acid Transporter 1 (4.3%,2)	L/I	Hyicolysin, Streptogrisin B	None	N/A
25	528-543	QDNEPEKPVADSETKM		A/Q	Procollagen I N-endopeptidase		
26	529-543	DNEPEKPVADSETKM		Q/D	Peptidyl-Asp metallo endopeptidase, Rice-tungro spherical virus-type endopeptidase		
27	531-543	EPEKPVADSETKM		N/E	Unknown		
28	191-211	AKAPAPAAPAAEPQAEAPVAS	Brain Acidic Membrane Protein (2,13%,2)	A/A	Saccharolysin, Peptidyl dipeptidase Dcp, Gametolysin, Peptidase T, **Bontoxilysin**, Glycyl aminopeptidase, Beta-Ala-His dipeptidase, Lysosomal dipeptidase I, Membrane Pro-X-carboxypeptidase, Pancreatic Elastase, Cytomegalovirus assemblin, Signal peptidase I, Dipeptidyl-peptidase II, C-terminal processing protease-1	S/S	Vibriolysin, Lysosomal dipeptidase I
29	191-220	AKAPAPAAPAAEPQAEAPVASSEQSVAVKE				None	N/A
1	2	3	4	5	6	7	8

Based on our observations of specific peptides during K^+-induced depolarization and the observations of others, the serine proteases appear to be particularly important in the ECF of the brain. Serine proteases, their natural inhibitors, serpins, and their receptors are found throughout the brain[101,102]. These enzymes contain a catalytic site consisting of a serine, histidine and aspartic acid residue. The exact proteases involved in the production of our 29 peptides remain unknown. However, co-release of PEA-derived peptides (e.g., Met- and Leu-enkephalin) and previously unknown fibrinopeptides (e.g., fibrinopeptide B_{1-13} and B_{2-14}) suggest that the serine proteases (e.g., PC1/3, PC2 and thrombin) are exceptionally important in the ECF of the brain.

Proteolytic Processing of PEA. The PEA-derived opioids, including Met- and Leu-enkephalin, are produced by the action of PC1/3, PC2 and numerous other proteases. PC1/3 and PC2 are calcium-dependent, serine proteases that are known to cleave PEA at mono- and di-basic sites in the Ca^{2+} rich, acidic environment of SV in the brain[72]. Differential processing of PEA by PC1/3 and PC2[103] corresponding with different distributions in neuronal tissue (e.g., striatum) and cell lines[104] has also been shown. Eight PEA-derived peptides, 6 novel peptides and Met- and Leu-enkephalin, were observed in the ECF of the brain by *in vivo* microdialysis-CLC-MS^2.

Seven of the 8 PEA-derived peptides that we observed support the general mechanism for proteolytic processing. Figure 4-2 shows the amino acid sequence of PEA and PEA-derived peptides discovered in this work. None of the 8 peptides observed originating from PEA (269 aa) were tryptic peptides (i.e., C-terminal R or K). This was expected as carboxypeptidase (e.g., carboxypeptidase C or H) cleavage of C-terminal mono- and di-basic amino acid residues following endopeptidase (e.g., PC1/3 and PC2) cleavage at mono- and di-basic sites is well-documented for PEA. Only 3 of the 8 peptides observed (i.e., peptide 1-YGGFM or Met-enkephalin, peptide 7-YGGFL or Leu-enkephalin and peptide 6-VGRPEWWMDYQ) can be explained solely by PC1/3 and PC2 cleavage. Four of the remaining 5 peptides (peptides 2-5 in Table 4-1) can be explained by PC1/3 and PC2 cleavage followed by carboxypeptidase cleavage.

However, peptide 8 cannot be explained simply by the action of the serine protease, acylaminoacyl peptidase, followed by carboxypeptidase cleavage (Table 4-1). Acylaminoacyl peptidase cleaves and acetylates the N-terminus of α-melaocyte stimulating hormone (α-MSH) in the brain, but we found peptide 8 to be unmodified. Therefore, another mechanism or unknown protease must produce peptide 8. Table 4-1 indicates that the metallopeptidase, lysosomal dipeptidase I, with specificity for cleaving SY bonds, is a good possibility.

MAQFLRLCIWLLALGSCLLATVQADCSQDCAKCSYRLVRPGDINFLACTLECEG
QLPSFKIWETCKDLLQVSKPEFPWDNIDMYKDSSKQDESHLLAKK**YGGFM**KR**Y**
GGFMKKMDELYPVEPEEEANGGEILAKR**YGGFM**KKDADEGDTLANSSDLLKEL
LGTGDNRAKD SHQQESTNND EDSTSKR**YGGFM**RGLKR *SPQLEDEAKELQ K*
R ***YGGFM*** *RRVG RPEWWMDYQKR* ***YGGFL*** KRFAESLPSDEEGESYSKEVPEME
KR**YGGFM**RF

Figure 4-2. Amino acid sequence of PEA and PEA-derived peptides discovered in this work. *In vivo* microdialysis-CLC-MS^2 revealed 6 previously unknown neuropeptides (underlined and boxed) and 2 known neuropeptides (bold). Peptide I (italics) was observed in BAM cells. Proteolytic processing of PEA produced peptides with 35% sequence coverage of the protein by amino acid count.

Extensive prior studies in brain tissue[45,105-108] and bovine chromaffin cells[79-94] have not revealed the presence of these 6 novel PEA-derived peptides. Because most previous work on proteolytic processing used antibodies to select peptides with certain

[109]epitopes, and because these peptides do not contain any homology with Met- or Leu-enkephalin, it is unlikely that these peptides would have been detected by any conventional processing studies. Figure 4-3 shows a non-exhaustive depiction of selected PEA-derived peptides by amino acid position. Unshaded peptides contain the sequence YGGFX where X is M or L while shaded peptides do not. Peptides observed in this work are indicated by an arrow and hypothetical intermediates (HIs) deduced from mono- and di-basic cleavage sites are indicated by a dashed arrow. Clearly, *in vivo* microdialysis-CLC-MS^2 has revealed a previously unknown subset of PEA-derived peptides without homology to Met- or Leu-enkephalin.

HIs for all of the PEA-derived peptides observed in this work are shown in Figures 4-3 and 4-4 for the 7 of 13 animals where PEA processing was observed (i.e., animals 1,2,4 and 10-13 in Table 3-1). We define a HI as the smallest precursor peptide (i.e., propeptide) containing all of the peptide sequences observed in each animal. Thus, HIs are relatively large peptides containing all of the peptide sequences observed in each animal. HIs are shown with and without (underlined) mono- or di-basic sites in Figures 4-3 and 4-4 for completeness. Predicted mono- and di-basic cleavage sites (i.e., K, R, KR, KK and RR), the peptides observed (bold and numbered by amino acid position according to Table 4-1) and selected other amino acids comprising the HI sequences (e.g., Q for peptides 4-7) are also shown[3]. Five HIs with mono- or di-basic sites and 3 without (underlined) were deduced from the peptide sequences observed in each animal (Figure 4-

---

[3] There are 6 copies of peptide 1-YGGFM (Met-enkephalin) in PEA and only 1 copy of peptide 7-YGGFL (Leu-enkephalin) as shown in Figure 4-2. We cannot determine which copy of Met-enkephalin was observed in the ECF by our approach; therefore, all amino acid positions are shown in Figures 4-3 and 4-4.

4). Spectral reproducibility studies *in vitro* and *in vivo* suggest that variability in the peptide sequences observed and the HIs predicted are, in fact, due to interanimal variability (Chapter3).

The HI for peptide 4-SPQLEDEAKE (peptide I_{1-10}), peptide 5-SPQLEDEAKELQ (peptide I_{1-12}), is produced in the ECF of the brain (Figure 4-4). Peptide 4-SPQLEDEAKE (peptide I_{1-10}) is an N-terminal cleavage product of peptide I, and it does not appear to have been produced from peptide I by bond cleavage of both the N- and C-terminus at mono- and di-basic sites that is typically associated with endopeptidase activity (e.g., YGGFM and YGGFL). The N-terminus appears to have been produced by cleavage at a dibasic site (i.e., KR); the C-terminus does not (i.e., LQ). Previously, we hypothesized that the production of peptide I_{1-10} occurs by endopeptidase (e.g. PC 1 and 2) cleavage of peptide I-SPQLEDEAKELQKRYGGFMRRVGRPEWWMDY, a known product of PEA-processing in purified BAM cells[26], to yield the intermediate SPQLEDEAKELQ which is then cleaved by a carboxypeptidase to yield the SPQLEDEAKE sequence detected *in vivo*. This pathway seemed likely, given the di-basic site in peptide I and the apparent ubiquity of exopeptidase activity in the brain extracellular space[27]. Measurement of peptide 5-SPQLEDEAKELQ (peptide I_{1-12}) in the ECF provides evidence for the proteolytic processing mechanism proposed above.

Other HIs may also be observed by *in vivo* microdialysis-CLC-MS^2. Relatively small HIs without mono- or di-basic sites (e.g., MDELYPVEPEEEANGEILA for peptides 2 and 3 and FAESLPSDEEGESYSKEVPEME for peptide 8) are amenable to sequencing by MS^2. However, many of the HIs with mono- or di-basic sites shown in Figures 4-3 and 4-4 (e.g., HI-11 or KKMDELEL2EEEANGEILAKR--------------------

KR4LQKR for peptide 2) are too large (e.g., 100aa for HI-ll) for peptide sequencing by low energy collision-induced dissociation (CID) in the QIT (although selected peptides may be quantifiable by time-segmented MS^2). Thus, our experiments revealed only a subset, albeit a previously unknown subset, of all possible bioactive peptides and intermediates in the ECF (see Chapter5 for future directions).

Comparison with Previous Work. Comparison of our results for PEA-processing in brain ECF with recent results for PEA-processing in BAM cells[43] and brain tissue[51] illustrates the importance of live animal studies for discrimination between intracellular peptides and those that are actually released into brain ECF. In the BAM cell study, preparative scale HPLC, antibodies for specific PEA-derived peptides, Edman sequencing and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOFMS) were combined to reveal 30 PEA-derived peptides[43]. In addition, several phosphorylated peptides were also found. All of the unmodified peptides that were discovered in that work are contained within Figure 4 (i.e., renumbered for the 30 aa signal peptide in PEA). The putative proteases cited for the production of these peptides included PC1/3, PC2, PAM, prohormone thiol-protease, propiomelanocortin-converting enzyme, serine proteases and carboxypeptidases. Five of the 30 peptides, corresponding to 3 novel cleavage sites (i.e., LA at aa positions 99-100, AK at 100-1001 and ME at 259-260) observed in the ECF during nicotine stimulation, were from the N- and C-terminal domain of PEA. In the brain tissue study, focused microwave irradiation and CLC-MS^2 were combined to reveal 550 endogenous peptides in the hypothalamus[51]. Also, several peptides with various PTMs were found. Nine unmodified PEA-derived peptides, corresponding to 3 novel cleavage sites (i.e., RS at

197-198, RV at 218-219 and YG at 263-264), spanned almost the entire domain of PEA. The putative proteases cited for production of these peptides included convertases, such as PC2, and carboxypeptidases. Interestingly, a PC2-derived peptide was also observed in that work.

None of the 30 PEA-derived peptides from the BAM cell study and only 4 of the 9 PEA-derived peptides from the brain tissue study (peptides 1,4,6 and 7) were observed in our experiments. Moreover, the 8 peptides we observed in brain ECF during K^+-stimulation, corresponding to 6 novel cleavage sites (i.e., LY at aa positions 117-118 for peptides 2 and 3, RS at 197-198 for peptides 4 and 5, RV at 218-219 for peptide 6 and SY at 251-252 for peptide 8), spanned almost the entire domain of PEA. These differences are not entirely surprising, given the different model systems and measurement techniques; however, many biological questions remain unanswered. For example, what is the relationship between 1) the peptides observed in the ECF of BAM cells 2) the peptides from brain tissue and 3) the peptides observed in the ECF of the brain following K^+-stimulation? Are the proteases and tertiary and quaternary structures of PEA the same? These and other biological questions merit further investigation.

Bioactivity of PEA-derived Peptides. The potential bioactivity of the 6 novel PEA-derived peptides that were discovered will require further study. Many of the BAM studies used trypsin followed by carboxypeptidase B cleavage to generate PEA-derived peptides for bioactivity assays such as the guinea pig ileum or opiate receptor binding assays. These and other PEA-derived peptides were found to be important in a large number of physiological and behavioral functions including pain and analgesia, immunity, appetite regulation and emotion, tolerance and dependence, mental illness and cardiovascular response[110]. The presence of peptides derived from PEA in the ECF is not surprising given the expression of this protein in the brain and the expectation that peptides "left-over" from the production of Met- and Leu-enkephalin by PEA-processing would be present in neuronal vesicles and co-released with peptides of known activity. However, it is unclear whether or not these other peptides possess bioactivity of their own. It has been suggested that PEA-derived peptides without bioactivity may direct bioactive peptides to specific target sites[111].

In an initial attempt to determine possible bioactivity of PEA-derived peptides, peptide $I_{1\text{-}10}$ was infused into the dialysis probe while simultaneously monitoring the levels of neuroactive amino acids such as γ-aminobutryric acid (GABA), glutamate, aspartate, glycine, and serine by *in vivo* microdialysis-capillary electrophoresis with laser-induced fluorescence detection ($n = 2$)[112]. An increase in aspartate and GABA was observed upon infusion of 10 μM peptide $I_{1\text{-}10}$ while glutamate levels remained the same (data not shown). GABA and its precursor, glutamate, are the major inhibitory and excitatory neurotransmitters in the brain, respectively. Consequently, the balance between these two is a subject of intense pharmacological investigation[113-118]. Previous microdialysis studies have shown that Leu-enkephalin activates opioid receptors and inhibits nicotine-stimulated GABA production[119]. In contrast, apomorphine stimulates GABA production while glutamate levels remain the same[120]. Receptor assays probing the binding of peptide $I_{1\text{-}10}$ to opioid receptors proved negative (data not shown). Therefore, activity of peptide $I_{1\text{-}10}$ is not mediated by the opioid receptors. Based on these results, we hypothesize that peptide $I_{1\text{-}10}$–stimulated GABA release is involved the regulation of other bioactive peptides such as Met- and Leu-enkephalin.

Proteolytic Processing of Neurogranin, Brain Acidic Membrane Protein and Excitatory Amino Acid Transporter 1. We observed peptides from the C-terminal domain of neurogranin (NG) [121-126], brain acidic membrane protein (BAMP)[99,127,128]and excitatory amino acid transporter 1 (EAAT1) [129] in the ECF of the brain. All of these proteins are substrates for protein kinase C (PKC) involved in synaptic plasticity and neurotransmitter release[121,129]. However, none of these proteins have any known peptide products. Seven novel peptides derived from NG, 2 from BAMP, and 4 from EAAT1 were discovered by *in vivo* microdialysis-CLC-MS^2.

We hypothesize that the peptides we observed from the C-terminal domains of these proteins are from extracellular domains that were exposed to proteases during depolarization. However, the orientation of NG remains unknown and the discovery that the C-terminal domain of BAMP resides on the external surface of SVs[99] does not support this theory. Moreover, the existing models for EAAT1[129,130], suggest that the C-terminal domain is intracellular, rather than extracellular, during depolarization-induced transporter-reversal[131,132]. Thus, no support for our hypothesis was provided by previous work on these proteins, and further investigations will be required.

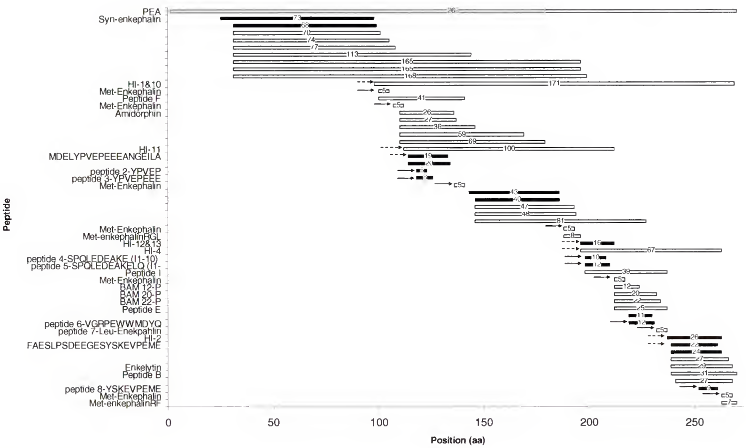


Figure 4-3. Non-exhaustive depiction of selected PEA-derived peptides by amino acid position. Unshaded peptides contain the sequence YGGFX where X is M or L while shaded peptides do not. The peptide length in amino acids is also shown. Hypothetical intermediates (HIs) with mono- or di-basic sites (listed by animal number) and without (SPQLEDEAKELQ, MDELYPVEPEEEANGEILA and FAESLPSDEEGESYSKEVPEME) are shown for the 7 of 13 animals where PEA processing was observed. Peptides observed in this work are indicated by an arrow and HIs are indicated by a dashed arrow.

Animal (n)	Hypothetical Intermediates of PEA Processing and Novel PEA-Processing Patterns Observed *In Vivo*
1	KK**1**KR**1**KK-----------------------------KR**1**KK--------KR**1**R-KR**4**LQKR**1**RR---KR**7**KRFAESLPSDEEGES**8**KR**1**R
2	KR**7**KRFAESLPSDEEGES**8**KR
4	KR**4**LQKR-RR**6**KR-KRFAESLPSDEEGES**8**KR
10	KK**1**KR**1**KKMDELEL **3** ANGEILAKR**1**KK--------KR**1**R-KR**4**LQKR**1**RR**6**KR**7**KRFAESLPSDEEGES**8**KR**1**R **5**
11	KKMDELEL**2**EEEANGEILAKR-------------------KR**4**LQKR
12	KR**4**LQKR
13	KR**4**LQKR

Figure 4-4. Hypothetical intermediates (HIs) of PEA processing and novel PEA-processing patterns observed *in vivo*. HIs with and without (underlined) mono- or di-basic sites for all of the PEA-derived peptides observed in this work are shown for the 7 of 13 animals where PEA processing was observed (i.e., animals 1,2,4 and 10-13 in Table 3-1). Predicted mono- and di-basic cleavage sites (i.e., K, R. KR, KK and RR), peptides observed (numbered by amino acid position according to Table 4-1) and selected other amino acids comprising the hypothetical intermediate sequences (e.g., Q for peptides 4-7) are shown. HIs are relatively large peptides containing all of the peptide sequences observed in each animal (not to scale).

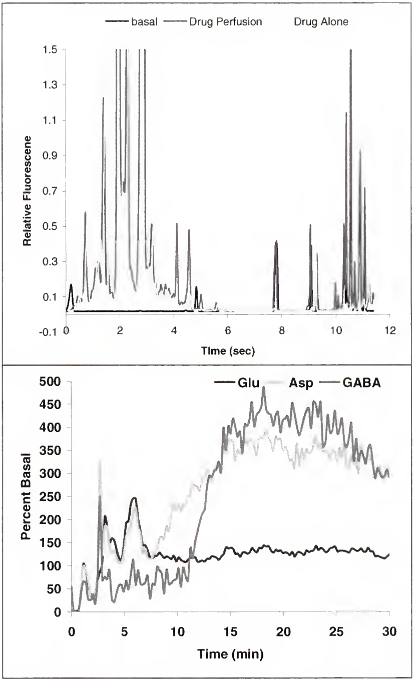


Figure 4-5. *In vivo* effect of peptide I_{1-10} on neurotransmitter levels. *In vivo* microdialysis-capillary electrophoresis with laser-induced fluorescence detection revealed a 3-4 fold increase in aspartate (Asp) and gamma-amino butyric acid (GABA) levels during perfusion of 10 μM peptide I_{1-10} through the microdialysis probe. An example electropherogram is shown in (A) and the temporal response of Asp, GABA, and Glu are shown in (B).

NG and BAMP belong to a small family of acidic, synaptic proteins that bind and sequester calmodulin (CaM)[99,125]. Neuronal depolarization leads to Ca^{2+} influx, phosphorylation of these proteins by PKC, release of CaM, and Ca^{2+}-CaM stimulated release of SVs[125]. Therefore, we hyopthesize that NG and BAMP are Ca^{2+} bound, unmodified and free of CaM during depolarization[125]. We did not observe any NG- or BAMP-derived peptides, phosphorylated or unmodified, containing known phosphorylation sites for PKC. However, the knowledge that the excitatory neurotransmitter glutamate is released from SVs during depolarization, combined with observations such as glutamate-stimulated NG phosphorylation[133] and enhanced NG phosphorylation during long-term potentiation[134], provide evidence in support of this hypothesis.

NG-, BAMP-, and EAAT1-derived peptides support the general mechanism for proteolytic processing. Five of the 7 NG-derived peptides (peptides 10-14 in Table 4-1) that we observed in the ECF of the brain can be explained either solely by PC2 cleavage or by PC2 cleavage followed by carboxypeptidase cleavage. The remaining 2 peptides (peptides 9 and 15 in Table 4-1) can be explained by various endopeptidases (e.g., the cysteine protease, cathepsin B) and by the serine protease dipeptidyldipeptidase II, respectively. The 2 BAMP-derived peptides (peptides 28 and 29 in Table 4-1) can be explained by various endopeptidases (e.g., the metallopeptidase, bontoxilysin) and, in the case of peptide 28, this is followed by carboxypeptidase cleavage. Three of the 4 EAAT1-derived peptides (peptides 24-26 in Table 4-1) can be explained by various endopeptidases (e.g., the serine protease, streptogrisin B). However, no putative proteases for peptide 27-EPEKPVADSETKM, cleaving the NE bond at the N-terminus, were found in the MEROPS database. This suggests that a novel protease, in addition to the novel cleavage site at aa positions 530-531, may have been discovered.

Decreased levels of NG (78 aa) have been observed during sleep-deprivation[135], aging[122], and brain injury studies[123] while very little is known about BAMP (220 aa). Various calcium-dependent cysteine proteases (e.g., calpain and caspase) released into the ECF during cell death may explain the decreased levels of NG observd in these earlier studies[136]. However, NG and BAMP are not known substrates for these proteases, and no evidence for the involvement of these proteases was found in this work.

EAAT1 (543 aa) is a Na^+-dependent glutamate/aspartate transport protein that mediates glutamate uptake mechanisms using a Na^+, K^+ electrochemical gradient as a driving force[129]. Glutamate uptake is enhanced during EAAT1 phosphorylation by PKC. We hypothesize that EAAT1 ceased to transport excess glutamate, possibly causing neurotoxic degeneration, disruption of the BBB and cell death. This is supported by greatly reduced extracellular Na^+ levels during K^+-induced depolarization.

Proteolytic Processing of Fibrinogen α and β. Fibrinogen (550 aa) is involved with thrombin in blood coagulation, and fibrinogen processing has been studied considerably in blood[137-139]. It is composed of 3 polypeptide chains (α, β and γ) and is proteolytically processed to fibrinopeptide A (α-derived) and fibrinopeptide B (β-derived) following binding to the protease thrombin. Fibrinopepides A and B assemble fibrin which coagulates blood by forming a polymeric clot. Fibrinogen-derived peptides are particularly important in diabetes[140] and cancer[141] research. Five novel peptides derived from fibrinogen α and 3 from fibrinogen β were discovered in this work.

Fibrinogen-derived peptides support the general mechanism for proteolytic processing. All of the 5 fibrinogen α-derived peptides (peptides 16-20 in Table 4-1) can be explained by various endopeptidases (e.g., the serine protease kallikrein). Two of the 3 fibrinogen β-derived peptides (peptides 21-22 in Table 4-1) can be explained by various endopeptidases (e.g., the serine protease thrombin that produces fibrinopeptide B). However, no putative proteases for peptide 23-TDSDKVDLSIAR, cleaving the TT bond at the N-terminus, were found in the MEROPS database. This suggests that a novel protease, in addition to the novel cleavage site at aa positions 530-531, may have been discovered.

Two of the 3 fibrinogen β-derived peptides, peptide 22-TTDSDKVDLSIA (fibrinopeptide B_{1-13}) and peptide 23- TDSDKVDLSIAR (fibrinopeptide B_{2-14}), are homologous with fibrinopeptide B (TTDSDKVDLSIAR), suggesting that depolarization caused disruption of the BBB and these peptides are involved in blood coagulation. While this remains a distinct possibility, the serine proteases have recently been shown to participate in various neuronal activities[102].

For example, proteolytically activated receptors (PARs), were found to be cleaved by thrombin[102,142,143]. PARs are activated by cleavage of an extracellular domain to form a new N-terminus which acts as a tethered ligand to bind to another extracellular domain for intracellular signaling. This phenomenon supports our hypothesis that the peptides we observed from the C-terminal domains of NG, BAMP and EAAT1 are from extracellular domains that were exposed to proteases during depolarization. It also suggests the possibility that these (e.g., fibrinogen) and other proteins may also be proteolytically activated.

Conclusion

The biological significance of the 29 peptide sequences and 6 protein precursors discovered by *in vivo* microdialysis-CLC-MS^2 in the ECF during K^+-induced depolarization remains unknown because the exact proteases involved and the mechanism that raised the levels of peptides in the ECF during K^+-induced depolarization remain unknown. However, examination of the peptide sequences as a function of the putative proteases involved provides important information. Depolarization caused increases in 8 PEA-derived peptides (corresponding to 6 novel peptides and 6 novel cleavage sites) and other peptides including neurogranin-, fibrinogen α and β- and excitatory amino acid transporter 1-derived peptides. Co-release of PEA-derived peptides (e.g., Met- and Leu-enkephalin) and previously unknown fibrinopeptides (e.g., fibrinopeptide B_{1-13} and B_{2-14}) suggest that the serine proteases (e.g., PC1/3, PC2 and thrombin) are exceptionally important in the ECF of the brain. Seven of the 8 PEA-derived peptides can be explained via cleavage by PC1/3 and PC2 followed by carboxypeptidase cleavage (e.g., carboxypeptidase H or C), supporting the general mechanism for proteolytic processing. Bioactivity of peptide I_{1-10}, a novel PEA-derived peptide, and co-release of known neuropeptides (i.e., the opioids Met- and Leu-enkephalin) suggests that important inter-neuron signaling mechanisms have been discovered.

CHAPTER 5
SUMMARY AND FUTURE DIRECTIONS

Summary

One of the most challenging problems in neurochemical monitoring *in vivo* is detecting neuropeptides. The difficulty stems from the low *in vivo* concentration of peptides (1-100 pM) which is ~10^6-fold lower than amino acid neurotransmitters. Thus, the detection method for peptides must be sensitive enough to quantify attomole levels in microliter sample volumes.

Analytical methods developed to address this challenge include LC- RIA[18], CLC-EC[21,22], CE-LIF[112] and MS^2 [24,54]. Of these techniques, MS^2 is the most promising because it is fast, gives sequence specific detection, and in principle can be used for any peptide. In ground-breaking work using this technique, endogenous Met-enkephalin[6] and neurotensin[7] were measured in 10 µL microdialysate fractions using an integrated preconcentration/desalting microelectrospray device, and the selected-reaction-monitoring (SRM) mode of a triple quadrupole mass spectrometer. While this method provided requisite sensitivity and selectivity, the lack of chromatographic separation and use of SRM precluded simultaneous monitoring of multiple peptides and required prior knowledge of the peptide for measurement. In addition, this method would not allow one of the most powerful aspects of MS^2 to be used, i.e.,identification of unknowns.

In the past 5 years, microdialysis-CLC-MS^2 has been applied to *in vivo* monitoring and discovering of endogenous peptides[9]. Several features of this method allowed improved results compared to the prior MS^2 work including: 1) miniaturized CLC columns and electrospray emitters[11,12] for better sensitivity, 2) mobile phase gradients optimized for sensitivity and chromatographic resolution, and 3) use of a quadrupole ion trap (QIT) which provides high sensitivity in full-scan MS^2[14] thus providing more information for peptide identification. Using this technique, multiple known and previously unknown neuropeptides were monitored simultaneously at low-attomole levels in 2 µL microdialysate fractions. The sensitivity was sufficient to allow 1 min temporal resolution; however, the on-line method was too slow to allow for such high temporal resolution and fractions were collected every 30 min.

An important feature of the system was the capability of identifying novel peptides by using the data-dependent MS^2 scan mode of the QIT and correlating raw MS^2 spectra to a database of protein sequences [16,17] (outline of the method is shown in Figure 1-7). In this way, unambiguous peptide sequences and protein precursors were assigned to the specific brain region sampled by the microdialysis probe. The proteomics tools used in this work were initially developed for unknown protein identification following controlled proteolytic cleavage (e.g., trypsin); however, the results of this study demonstrated that it may be useful for discovering previously unidentified peptides following unknown proteolytic processing of protein precursors in the brain.

We found that peptide sequencing and protein precursor identification at attomole levels necessitates sequence validation. Many of the MS^2 spectra collected at attomole levels in the data-dependent MS^2 scan mode corresponded to peptide sequences with database-searching scores below the accepted limits for unambiguous protein precursor identification. The false-positive rate for database-derived, non-tryptic peptide sequences has not been rigorously evaluated. Therefore, non-tryptic peptide sequences must be validated before they are reported.

Chapter 2 described the development, evaluation and application of a novel *in vivo* microdialysis-CLCMS^2 system for monitoring peptides. Simultaneous, on-line monitoring of Met-enkephalin, Leu-enkephalin and a previously unknown peptide at attomole levels with 30-min temporal resolution were described. Preliminary evaluation of the system for discovering previously unknown peptides by data-dependent MS^2 was also investigated. Validation of the Sequest[55]-derived sequence for a previously unknown peptide, peptide I_{1-10} (SPQLEDEAKE), was obtained by matching the MS^2 and MS^3 spectra of the synthetic peptide SPQLEDEAKE with MS^2 and MS^3 spectra collected *in vivo*. While this information, combined with matching chromatographic retention times, provided confident validation of the peptide discovered *in vivo*, it would be time-consuming and expensive for a large number of peptides.

Chapter 3 described the development, evaluation and application of a proteomics-based data-reduction strategy in conjunction with the *in vivo* microdialysis-CLC-data-dependent-MS^2 system for discovering previously unknown peptides. The reproducibility of peptide sequencing and protein precursor identification at attomole levels was also investigated. Increasingly stringent criteria were applied at successive data-reduction steps in order to improve our confidence in the peptide sequences discovered *in vivo*. In step 1, a threshold was applied to keep only MS^2 spectra with a total ion current greater than 5e + 5. We found that similar numbers of MS^2 spectra were found *in vitro* and *in vivo* suggesting that the number of MS^2 spectra collected was primarily concentration-dependent. In step 2, subtractive analysis was performed to keep MS^2 spectra observed only during K$^+$-induced depolarization conditions. We found that no validated MS^2 spectra were observed at basal conditions; therefore, MS^2 spectra of contaminants (e.g., from the microdialysis probe) were eliminated prior to database searching with Sequest. In step 3, a threshold was applied to keep only significant MS^2 spectra (i.e., with a Sequest ΔCn score greater than 0.1). We found that similar fractions of significant MS^2 spectra were found *in vitro* and *in vivo* suggesting that the fraction of significant MS^2 spectra collected was also concentration-dependent. In step 4, electronic validation of Sequest-derived sequences was obtained by matching the Sequest- and Mascot[59]-derived sequences. We found this to be more stringent than validation by requiring two peptide sequences per protein precursor. We also found that Sequest provided more information, while Mascot provided less false-positives. In step 5, additional electronic validation was provided investigated by the *de novo* sequencing program Lutefisk[60]. We found that Lutefisk further validated approximately half the sequences validated by database-searching. In step 6, incorrect sequence assignments by database-searching programs were found by searching significant MS^2 spectra against a user library[63,144,145] of electronically validated MS^2 spectra. We found this 'rescued' a few incorrectly assigned MS^2 spectra and thus slightly improved the reproducibility of peptide sequencing and protein precursor identification *in vivo*.

These studies provided proof-of-principle for our data-reduction strategy and demonstrated that *in vivo* microdialysis-CLC-data-dependent-MS^2 is an effective means of exploring the brain. Indeed, 27 previously unknown peptides and 2 known neuropeptides were sequenced by this approach These experiments also revealed that only one peptide sequence per protein precursor is necessary for confident protein identification if the peptide sequence is sufficiently validated. Comparing sequences obtained from two different database-searching programs (the 2nd more stringent than the

1^{st}) and additional validation by one *de novo* sequencing program was our preferred method of sequence validation *in lieu* of peptide synthesis or manual sequence interpretation.

Chapter 4 described the biological significance of the 29 peptide sequences and 6 protein precursors discovered by *in vivo* microdialysis-CLC-MS^2 in the ECF during K$^+$-induced depolarization. The exact proteases involved and the mechanism that raised the levels of peptides in the ECF during K$^+$-induced depolarization remain unknown. However, examination of the peptide sequences as a function of the putative proteases involved provided important information. Depolarization caused increases in 8 PEA-derived peptides (corresponding to 6 novel peptides and 6 novel cleavage sites) and other peptides including neurogranin-, fibrinogen α and β- and excitatory amino acid transporter 1-derived peptides. Co-release of PEA-derived peptides (e.g., Met- and Leu-enkephalin) and previously unknown fibrinopeptides (e.g., fibrinopeptide B_{1-13} and B_{2-14}) suggests that the serine proteases (e.g., PC1/3, PC2 and thrombin) are exceptionally important in the ECF of the brain. Seven of the 8 PEA-derived peptides can be explained via cleavage by PC1/3 and PC2 followed by carboxypeptidase cleavage (e.g., carboxypeptidase H or C), supporting the general mechanism for proteolytic processing. Bioactivity of peptide I_{1-10}, a novel PEA-derived peptide, and co-release of known neuropeptides (i.e., the opioids Met- and Leu-enkephalin) suggests that important inter-neuron signaling mechanisms were discovered.

However, *in vivo* microdialysis-CLC-data-dependent-MS^2, like all techniques, suffers from some limitations. The most serious limitation of this approach is that the *in vivo* results described herein required unknown protease specificity to be selected when running database-searching or *de novo* sequencing programs. Unknown protease specificity, combined with variable post-translational modifications (PTMs), significantly increases processing time. Thus, even a 2 GHz computer with 1 GB of RAM may require days to evaluate the MS^2 spectra collected during a single CLC-MS^2 analysis! Another serious limitation is that the technique is limited to relatively small peptides due to poor fragmentation of larger peptides by the mass spectrometer and poor recovery of larger peptides by the microdialysis probe. This makes studying processing of endogenous neuropeptides problematic because gaps in our knowledge of peptide intermediates, as described in chapter 4, will remain open unless this problem is resolved. Lastly, multi-dimensional techniques such as ours are technically difficult since each dimension must be optimized in order to insure full-functionality of the system. Troubleshooting can be nightmarish if the problem isn't systematically investigated, and sharing the system with multiple users of varying experience further complicates matters. However, we have determined the symptoms, diagnosis and treatment for the most common problems in each dimension of the technique as described in appendix C.

Future Directions

In principle, this method could be used to identify localized neurochemical changes that occur in response to any physiological or behavioral state. Our laboratory explored this possibility by comparing *in vivo* microdialysis-CLC-MS^2 data obtained from animals during sleep and prolonged wakefulness as shown in Figure 5-1[146]. While the detected compounds have yet to be identified, the data clearly illustrate a differential peptide profile in the extracellular space under these two conditions. Detection of these chemical changes could be used to identify compounds that control the sleep state. One can envision many applications for identifying biomarkers and/or signaling molecules of behaviors or physiological states as the method is refined.

Continuing advances in proteomics may enable more detailed analysis of dialysates and similar samples in the future. For example, preparative 2-dimensional electrophoresis of human cerebral spinal fluid followed by trypsin proteolysis and matrix-assisted laser desorption/ionization-time of flight mass spectrometry (MALDI-TOFMS) revealed 21 proteins not previously observed by traditional methods[147] and CLC-MS^2 has allowed identification of over 1500 peptides in brain tissue samples[148]. These techniques offer complementary sequence information for *in vivo* microdialysis-CLC-MS^2 studies (from peptides that ionize well by MALDI and from peptides found in brain tissue, respectively). Another interesting approach is the detection of neuropeptides in carboxypeptidase E deficient (i.e., 'knockout') mice. The lack of this protease results in elevated levels of neuropeptides which facilitated identification of many novel peptides in tissue samples by CLC-MS^2 [149]. The combined use of CE with MALDI-MS to analyze dialysates[47] is also intriguing. While the technique is not yet sensitive enough for detecting endogenous peptides, it has been demonstrated that addition of exogenous peptide to the dialysis probe allows observation of *in vivo* processing of the peptide.

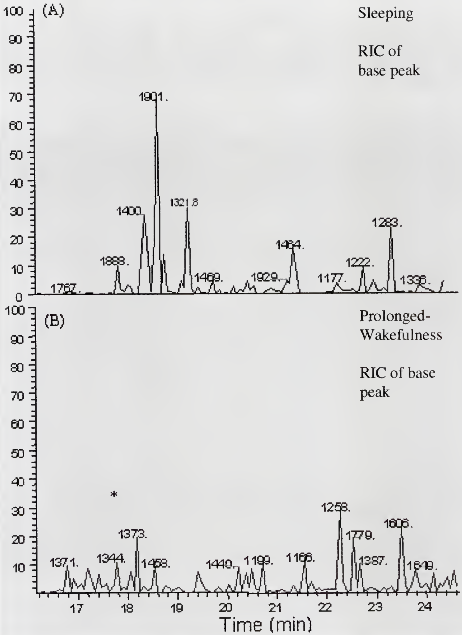


Figure 5-1. *In vivo* microdialysis-CLC-MS^2 base peak reconstructed ion chromatograms (RICs) show the most abundant product ions observed during sleeping (A) and prolonged-wakefulness (B) states. The MS^2 spectrum for the doubly charged unknown precursor ion at *m/z* 1282 and its most abundant product ion at *m/z* 1373 is labeled with an asterisk (B). The MS^2 spectrum for the unknown precursor ion at *m/z* 1282 was observed only in the forced-wakefulness state (B). The microdialysis probe was implanted in the hypothalamus of a male rat.

Recent advances in 'top-down' protein sequencing by electron capture dissociation (ECD) in a Fourier-transform ion cyclotron resonance mass spectrometer (FT-ICRMS)[150] suggest that larger peptides will soon be accessible by CLC-MS^2. However, this poses even greater computational problems for peptide sequencing and protein precursor identification by database searching. There are no elegant computational solutions at this time, although probability-based scoring systems such as Mascot significantly improve confidence in the results[58,59,151-156]. In order to simultaneously characterize all the products of proteolytic processing by *in vivo* microdialysis-CLC-MS^2, concomitant advances in the recovery of larger peptides through the microdialysis probe will also be required. If achieved, then rigorous determination of the order of appearance of all peptide intermediates could be performed. This would provide sequence-specific kinetic information, limited by the detection limit and temporal resolution of CLC-MS^2, on the proteases and peptides (i.e., substrates and products) important in the ECF of living animals. This information is critical for investigations of cell-signaling and for the production of recombinant proteins. Despite the limitations of *in vivo* microdialysis-CLC-MS^2, it is evident that this technique, in conjunction with proteomic methods, has growing potential for monitoring and discovering peptides in living animals and promising applications in biotechnology.

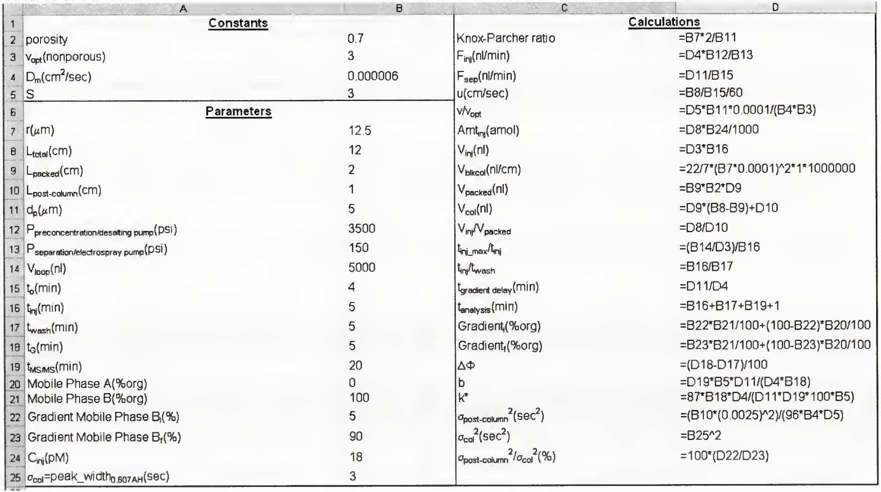
APPENDIX A
EXAMPLE OF CLC-MS^2 CALCULATIONS

Table A-1 is an example of the CLC-MS^2 calculations used throughout this work. Important constants and parameters are shown for CLC-MS^2 calculations during typical operating conditions. This example is non-exhaustive and is only intended as a guide, in conjunction with selected references[6,9], to those who may be pursuing similar lines of research

Table A-1. Example of CLC-MS^2 calculations.

Constants		**Calculations**	
porosity	0.7	Knox-Parcher ratio	5
v_{opt}(nonporous)	3	F_{inj}(nl/min)	327
D_m(cm^2/sec)	6.00E-06	F_{sep}(nl/min)	14
S	3	u(cm/sec)	0.05
Parameters		v/v_{opt}	1.4
r(πm)	12.5	Amt_{inj}(amol)	29
L_{total}(cm)	12	V_{inj}(nl)	1633
L_{packed}(cm)	2	V_{blkcol}(nl/cm)	4.9
$L_{post\text{-}column}$(cm)	1	V_{packed}(nl)	6.9
d_p(πm)	5	V_{col}(nl)	56
$P_{preconcentration/desalting\ pump}$(psi)	3500	V_{inj}/V_{packed}	238
$P_{separation/electrospray\ pump}$(psi)	150	t_{inj_max}/t_{inj}	3
V_{loop}(nl)	5000	t_{inj}/t_{wash}	1
t_o(min)	4.0	$t_{gradient\ delay}$(min)	4.0
t_{inj}(min)	5	$t_{analysis}$(min)	31
t_{wash}(min)	5	$Gradient_i$(%org)	5
t_G(min)	5	$Gradient_f$(%org)	90
$t_{MS/MS}$(min)	20	$\pi\{$	0.85
Mobile Phase A(%org)	0	b	2.04
Mobile Phase B(%org)	100	k*	0.43
Gradient Mobile Phase B_i(%)	5	$\pi_{post\text{-}column}{}^2$(sec^2)	0.22
Gradient Mobile Phase B_f(%)	90	$\pi_{col}{}^2$(sec^2)	9.00
C_{inj}(pM)	18	$\pi_{post\text{-}column}{}^2/\pi_{col}{}^2$(%)	2.41
π_{col}=peak_width$_{0.607AH}$(sec)	3		

Table A-1. Continued.



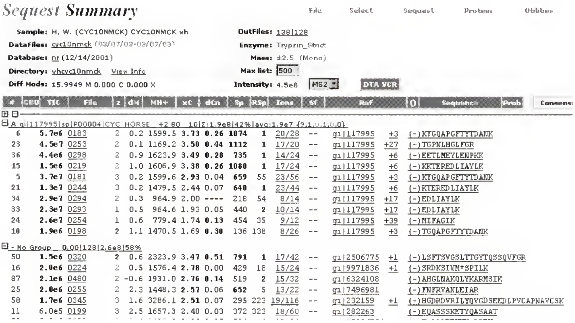
	A	B	C	D
1	**Constants**		**Calculations**	
2	porosity	0.7	Knox-Parcher ratio	=B7*2/B11
3	v_{opt}(nonporous)	3	F_{inj}(nl/min)	=D4*B12/B13
4	D_m(cm^2/sec)	0.000006	F_{sep}(nl/min)	=D11/B15
5	S	3	u(cm/sec)	=B8/B15/60
6	**Parameters**		v/v_{opt}	=D5*B11*0.0001/(B4*B3)
7	r(μm)	12.5	Amt_{inj}(amol)	=D8*B24/1000
8	L_{total}(cm)	12	V_{inj}(nl)	=D3*B16
9	L_{packed}(cm)	2	V_{blkcol}(nl/cm)	=22/7*(B7*0.0001)^2*1*1000000
10	$L_{post\text{-}column}$(cm)	1	V_{packed}(nl)	=B9*B2*D9
11	d_p(μm)	5	V_{col}(nl)	=D9*(B8-B9)+D10
12	$P_{preconcentration/desalting\ pump}$(psi)	3500	V_{inj}/V_{packed}	=D8/D10
13	$P_{separation/electrospray\ pump}$(psi)	150	t_{inj_max}/t_{inj}	=(B14/D3)/B16
14	V_{loop}(nl)	5000	t_{inj}/t_{wash}	=B16/B17
15	t_o(min)	4	$t_{gradient\ delay}$(min)	=D11/D4
16	t_{inj}(min)	5	$t_{analysis}$(min)	=B16+B17+B19+1
17	t_{wash}(min)	5	$Gradient_i$(%org)	=B22*B21/100+(100-B22)*B20/100
18	t_G(min)	5	$Gradient_f$(%org)	=B23*B21/100+(100-B23)*B20/100
19	$t_{MS/MS}$(min)	20	$\Delta\Phi$	=(D18-D17)/100
20	Mobile Phase A(%org)	0	b	=D19*B5*D11/(D4*B18)
21	Mobile Phase B(%org)	100	k*	=87*B18*D4/(D11*D19*100*B5)
22	Gradient Mobile Phase B_i(%)	5	$\sigma_{post\text{-}column}{}^2$(sec^2)	=(B10*(0.0025)^2)/(96*B4*D5)
23	Gradient Mobile Phase B_f(%)	90	$\sigma_{col}{}^2$(sec^2)	=B25^2
24	C_{inj}(pM)	18	$\sigma_{post\text{-}column}{}^2/\sigma_{col}{}^2$(%)	=100*(D22/D23)
25	σ_{col}=peak_width$_{0.607AH}$(sec)	3		

APPENDIX B

EXAMPLE OF PEPTIDE SEQUENCING AND PROTEIN PRECURSOR IDENTIFICATION

This example is non-exhaustive and is only intended as a guide, in conjunction with selected references[6,55,58-60,63,144,145,157], to those who may be pursuing similar lines of research In chapter 5, a summary of this work and future directions were presented. I strongly suggest investigating newer approaches to peptide sequencing and protein precursor identification in conjunction with the methods described here because the field of proteomics is developing rapidly!

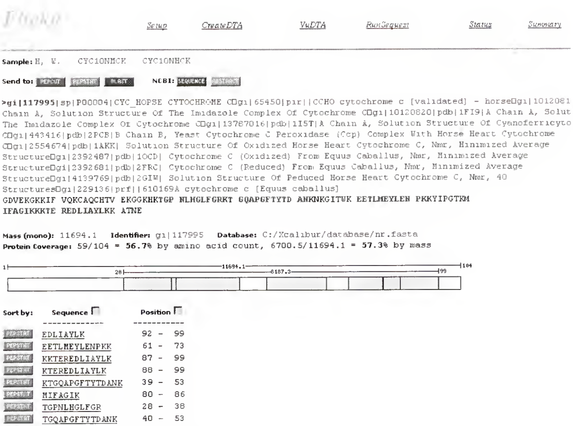
The following figures illustrate the electronic validation approach described in chapter 3 for a tryptic digest of cytochrome c. Database-searching should be performed on Sequest 1st and Mascot 2nd as Sequest is significantly faster. However, PTMs are more easily investigated with Mascot (i.e., using the error tolerant search function) because Sequest is currently limited to only 3 variable PTMs per search. Excellent instructions for Xcalibur (including the creation of user-libraries), Sequest, Mascot and Lutefisk are freely available from the respective manufacturers. In addition, the manufacturer of Mascot maintains a website a useful collection of internet hyperlinks to various proteomic resources (www.matrixscience.com). I suggest running all peptide sequencing and protein precursor identification programs on the fastest dual processor personal computer available and dedicating this computer to CLC-MS^2 analysis. Sequest, Mascot and Lutefisk should be installed on the c: drive to preserve the default file configurations. Protein databases should be downloaded from NCBI and installed into both Sequest and Mascot to enable direct comparisons and to avoid confusion resulting from differences in protein nomenclature. Whenever possible, I suggest searching the species-specific databases instead of the nonredundant databases in order to reduce processing time. Lastly, all protein databases, including the 'nonredundant' databases, are redundant (i.e., there are multiple entries for many proteins); therefore, it is far simpler to think in terms of peptide sequences rather than protein precursors when correlating results from different CLC-MS^2 datasets.



Sample: H, W. (CYC10NMCK) CYC10NMCK wh
DataFiles: cyc10nmck (03/07/03-03/07/03)
Database: nr (12/14/2001)
Directory: whcyc10nmck View Info
Diff Mods: 15.9949 M 0.000 C 0.000 X
OutFiles: 138|128
Enzyme: Trypsin_Strict
Mass: ±2.5 (Mono)
Max list: 500
Intensity: 4.5e8

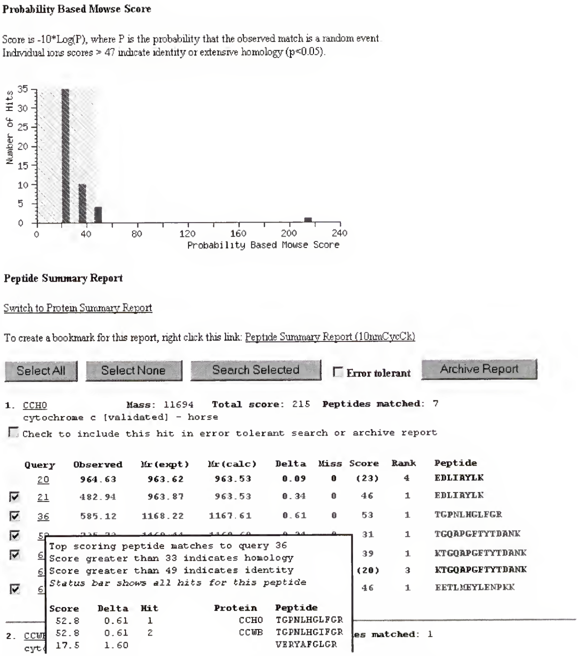
#	CBU	TIC	File	z	dM	MH+	xC	dCn	Sp	RSp	Ions	Sf	Ref	()	Sequence	Prob
A gi|117995|sp|P00004|CYC_HORSE_ +2.80 10|Σ:1.9e8|42%|avg:1.9e7 {9,1,0,1,0,0}																
6		5.7e6	0183	2	0.2	1599.5	3.73	0.26	1074	1	20/28	--	gi|117995	+3	(-)KTGQAPGFTYTDANK	
23		4.5e7	0253	2	0.1	1169.2	3.50	0.44	1112	1	17/20	--	gi|117995	+27	(-)TGPNLHGLFGR	
36		4.4e6	0298	2	0.9	1623.9	3.49	0.28	735	1	14/24	--	gi|117995	+6	(-)EETLMEYLENPKK	
15		1.5e6	0219	2	1.0	1606.9	3.38	0.26	1080	1	17/24	--	gi|117995	+6	(-)KKTEREDLIAYLK	
5		3.7e7	0181	3	0.2	1599.6	2.93	0.04	659	55	23/56	--	gi|117995	+3	(-)KTGQAPGFTYTDANK	
21		1.3e7	0244	3	0.2	1479.5	2.44	0.07	640	1	23/44	--	gi|117995	+6	(-)KTEREDLIAYLK	
34		2.9e7	0294	2	0.3	964.9	2.00	----	218	54	8/14	--	gi|117995	+17	(-)EDLIAYLK	
33		2.3e7	0293	1	0.5	964.6	1.93	0.05	440	2	10/14	--	gi|117995	+17	(-)EDLIAYLK	
24		2.6e7	0254	1	0.6	779.4	1.74	0.13	454	35	9/12	--	gi|117995	+39	(-)MIFAGIK	
10		1.9e6	0198	2	1.1	1470.5	1.69	0.30	136	138	8/26	--	gi|117995	+3	(-)TGQAPGFTYTDANK	
- No Group 0.00|128|2.6e8|58%																
50		1.5e6	0320	2	0.6	2323.9	3.47	0.51	791	1	17/42	--	gi|2506775	+1	(-)LSFTSVGSLTTGYTQSSQVFGR	
16		2.0e6	0224	2	0.5	1576.4	2.78	0.00	429	18	15/24	--	gi|9971836	+1	(-)SRDKSIWM*SPILK	
87		2.1e6	0480	2	-0.6	1931.0	2.76	0.14	519	2	15/32	--	gi|6324108		(-)AHGLNAKQLYKARMSIK	
25		2.0e6	0255	2	2.3	1448.3	2.57	0.06	652	5	13/22	--	gi|7496981		(-)FNFRVANLEIAR	
58		1.7e6	0345	3	1.6	3286.1	2.51	0.07	295	223	19/116	--	gi|232159	+1	(-)HGDRDVRILYQVGDSEEDLPVCAPNAVCSK	
11		6.0e5	0199	3	2.5	1657.3	2.40	0.03	372	323	18/60	--	gi|282263		(-)EQASSSKETYQASAAT	

Figure B-1. Sequest summary report for a tryptic digest of cytochrome c. I have found that a 10 fmol / μL (nM) aqueous solution of cytochrome c digested with trypsin (1:100 dilution of stock obtained from Michrom Bioresources, Auburn, CA) is a good starting point for beginners (i.e., this is 10-100 fold more concentrated than the peptides discovered in this work). A 5 μL injection of this sample onto the CLC column corresponds to ~ 50 fmol of each tryptic peptide injected on-column (i.e., 104 amino acids / 10 = 10 tryptic peptides and (10 fmol / μL)(5 μL) = 50 fmol). Significant MS^2 spectra (rows with bold Xcorr or ΔCn scores) were revealed by searching the non-redundant protein database using strict trypsin specificity and the default precursor and product ion mass tolerances. The peptide TGPNLHGLFGR received a significant Xcorr score of 3.50 and a significant ΔCn score of 0.44 in Sequest.



```
Sample: H, W.    CYC1ONMCK    CYC1ONHCK

Send to:                NCBI:

>gi|117995|sp|P00004|CYC_HOPSE CYTOCHROME C□gi|65450|pir||CCHO cytochrome c [validated] - horse□gi|1012081
Chain A, Solution Structure Of The Imidazole Complex Of Cytochrome C□gi|10120820|pdb|1FI9|A Chain A, Solut
The Imidazole Complex Of Cytochrome C□gi|13787016|pdb|1I5T|A Chain A, Solution Structure Of Cyanoferricyto
C□gi|443416|pdb|2PCB|B Chain B, Yeast Cytochrome C Peroxidase (Ccp) Complex With Horse Heart Cytochrome
C□gi|2554674|pdb|1AKK| Solution Structure Of Oxidized Horse Heart Cytochrome C, Nmr, Minimized Average
Structure□gi|2392487|pdb|1OCD| Cytochrome C (Oxidized) From Equus Caballus, Nmr, Minimized Average
Structure□gi|2392681|pdb|2FRC| Cytochrome C (Reduced) From Equus Caballus, Nmr, Minimized Average
Structure□gi|4139769|pdb|2GIW| Solution Structure Of Peduced Horse Heart Cytochrome C, Nmr, 40
Structures□gi|229136|prf||610169A cytochrome c [Equus caballus]
GDVEKGKKIF VQKCAQCHTV EKGGKHKTGP NLHGLFGRKT GQAPGFTYTD ANKNKGITWK EETLMEYLEN PKKYIPGTKM
IFAGIKKKTE REDLIAYLKK ATNE

Mass (mono): 11694.1   Identifier: gi|117995   Database: C:/Xcalibur/database/nr.fasta
Protein Coverage: 59/104 = 56.7% by amino acid count, 6700.5/11694.1 = 57.3% by mass

Sort by:    Sequence ☐        Position ☐
            -------------     -----------
            EDLIAYLK           92 -  99
            EETLMEYLENPKK      61 -  73
            KKTEREDLIAYLK      87 -  99
            KTEREDLIAYLK       88 -  99
            KTGQAPGFTYTDANK    39 -  53
            MIFAGIK            80 -  86
            TGPNLHGLFGR        28 -  38
            TGQAPGFTYTDANK     40 -  53
```

Figure B-2. Sequest cytochrome c report for a tryptic digest of cytochrome c. Eight tryptic peptides correspond to sequence coverage of 57% by amino acid count. Peptide TGPNLHGLFGR spans amino acid positions 28-38 in cytochrome c.



Probability Based Mowse Score

Score is -10*Log(P), where P is the probability that the observed match is a random event.
Individual ions scores > 47 indicate identity or extensive homology (p<0.05).

Peptide Summary Report

Switch to Protein Summary Report

To create a bookmark for this report, right click this link: Peptide Summary Report (10nmCycCk)

1. CCHO **Mass:** 11694 **Total score:** 215 **Peptides matched:** 7
cytochrome c [validated] - horse

☐ Check to include this hit in error tolerant search or archive report

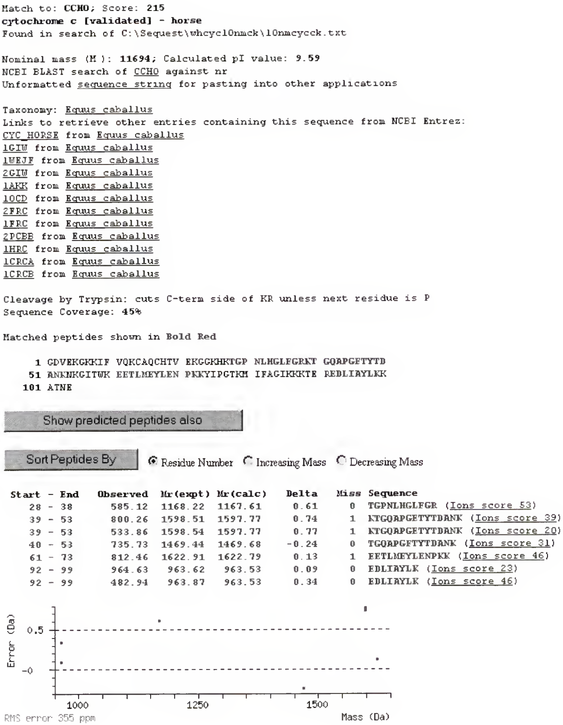
	Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Rank	Peptide
	20	**964.63**	**963.62**	**963.53**	**0.09**	**0**	**(23)**	**4**	**EDLIAYLK**
☑	21	482.94	963.87	963.53	0.34	0	46	1	EDLIAYLK
☑	36	585.12	1168.22	1167.61	0.61	0	53	1	TGPNLHGLFGR
☑	5						31	1	TGQAPGFTYTDANK
☑	6						39	1	KTGQAPGFTYTDANK
	6						**(20)**	**3**	**KTGQAPGFTYTDANK**
☑	6						46	1	EETLMEYLENPKK

Top scoring peptide matches to query 36
Score greater than 33 indicates homology
Score greater than 49 indicates identity
Status bar shows all hits for this peptide

Score	Delta	Hit	Protein	Peptide
52.8	0.61	1	CCHO	TGPNLHGLFGR
52.8	0.61	2	CCWB	TGPNLHGIFGR
17.5	1.60			VERYAFGLGR

2. CCW **es matched:** 1
cyt

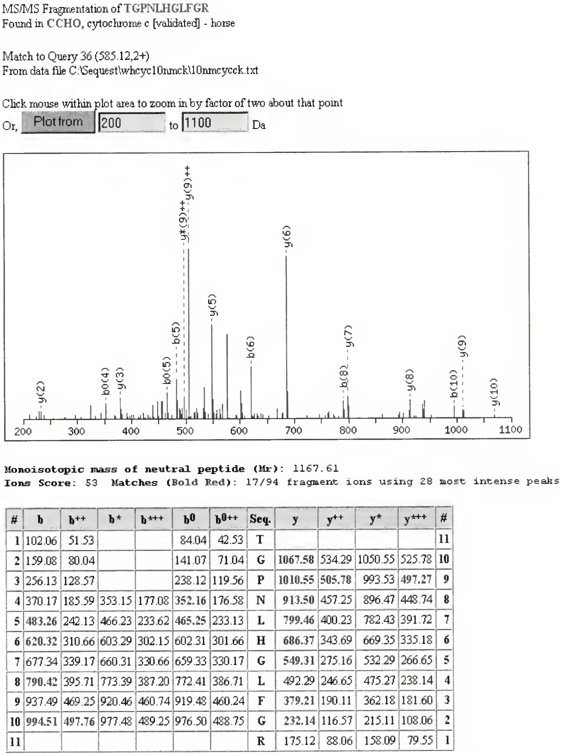
Figure B-3. Mascot summary report for a tryptic digest of cytochrome c. The peptide TGPNLHGLFGR received a score of 53 in Mascot indicating at least homology with 95% probability. In this case, a score 33 indicates homology and a score of 49 indicates identity. This provides validation for the Sequest-derived peptide sequence and protein precursor identification.



```
Match to: CCHO; Score: 215
cytochrome c [validated] - horse
Found in search of C:\Sequest\whcyc10nmck\10nmcyck.txt

Nominal mass (M ): 11694; Calculated pI value: 9.59
NCBI BLAST search of CCHO against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Equus caballus
Links to retrieve other entries containing this sequence from NCBI Entrez:
CYC_HORSE from Equus caballus
1GIW from Equus caballus
1WEJF from Equus caballus
2GIW from Equus caballus
1AKK from Equus caballus
1OCD from Equus caballus
2FRC from Equus caballus
1FRC from Equus caballus
2PCBB from Equus caballus
1HRC from Equus caballus
1CRCA from Equus caballus
1CRCB from Equus caballus

Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: 45%

Matched peptides shown in Bold Red

     1 GDVEKGKKIF VQKCAQCHTV EKGGKHKTGP NLHGLFGRKT GQAPGFTYTD
    51 ANKNKGITWK EETLMEYLEN PKKYIPGTKM IFAGIKKKTE REDLIAYLKK
   101 ATNE
```

```
 Start - End    Observed  Mr(expt) Mr(calc)   Delta   Miss Sequence
    28 - 38       585.12  1168.22  1167.61     0.61     0  TGPNLHGLFGR (Ions score 53)
    39 - 53       800.26  1598.51  1597.77     0.74     1  KTGQAPGFTYTDANK (Ions score 39)
    39 - 53       533.86  1598.54  1597.77     0.77     1  KTGQAPGFTYTDANK (Ions score 20)
    40 - 53       735.73  1469.44  1469.68    -0.24     0  TGQAPGFTYTDANK (Ions score 31)
    61 - 73       812.46  1622.91  1622.79     0.13     1  EETLMEYLENPKK (Ions score 46)
    92 - 99       964.63   963.62   963.53     0.09     0  EDLIAYLK (Ions score 23)
    92 - 99       482.94   963.87   963.53     0.34     0  EDLIAYLK (Ions score 46)
```

Figure B-4. Mascot cytochrome c report for a tryptic digest of cytochrome c. Seven tryptic peptides correspond to sequence coverage of 45% by amino acid count. Peptide TGPNLHGLFGR spans amino acid positions 28-38 in cytochrome c. The 2 peptides spanning positions 87-99 and 88-99 in Sequest only span positions 92-99 in Mascot. Thus, these peptides were not validated by Mascot. In general, Mascot provides less false-positives and less information. Thus, it is a valuable check of the Sequest results.



MS/MS Fragmentation of TGPNLHGLFGR
Found in CCHO, cytochrome c [validated] - horse

Match to Query 36 (585.12,2+)
From data file C:\Sequest\whcyc10nmck\10nmcycck.txt

Click mouse within plot area to zoom in by factor of two about that point
Or, Plot from 200 to 1100 Da

Monoisotopic mass of neutral peptide (Mr): 1167.61
Ions Score: 53 **Matches (Bold Red):** 17/94 fragment ions using 28 most intense peaks

#	b	b^{++}	b^{*}	b^{*++}	b^{0}	b^{0++}	Seq.	y	y^{++}	y^{*}	y^{*++}	#
1	102.06	51.53			84.04	42.53	T					11
2	159.08	80.04			141.07	71.04	G	**1067.58**	534.29	1050.55	525.78	10
3	256.13	128.57			238.12	119.56	P	**1010.55**	**505.78**	993.53	**497.27**	9
4	370.17	185.59	353.15	177.08	**352.16**	176.58	N	**913.50**	457.25	896.47	448.74	8
5	**483.26**	242.13	466.23	233.62	**465.25**	233.13	L	**799.46**	400.23	782.43	391.72	7
6	**620.32**	310.66	603.29	302.15	602.31	301.66	H	**686.37**	343.69	669.35	335.18	6
7	677.34	339.17	660.31	330.66	659.33	330.17	G	**549.31**	275.16	532.29	266.65	5
8	**790.42**	395.71	773.39	387.20	772.41	386.71	L	492.29	246.65	475.27	238.14	4
9	937.49	469.25	920.46	460.74	919.48	460.24	F	**379.21**	190.11	362.18	181.60	3
10	**994.51**	**497.76**	977.48	489.25	976.50	488.75	G	**232.14**	116.57	215.11	108.06	2
11							R	175.12	88.06	158.09	79.55	1

Figure B-5. Mascot annotated report for the tryptic peptide TGPNLHGLFGR from the tryptic digest of cytochrome c. An overlapping partial b-type ion series and almost complete y-type ion series are indicative of a significant MS^2 spectrum.

```
Lutefisk1900 v1.3.2
Copyright 1996-2001 Richard S. Johnson
Run Date: Mon Mar 24 07:52:04 2003
 Filename: cyc10nmck.0253.0253.2.dta
 Molecular Weight: 1168.22  Molecular Weight Tolerance:  2.00  Fragment Ion
Tolerance: 0.80
 Ion Offset: 0.00 Charge State: 2 Centroided or Pre-processed Data
 ??? Digest    Tryptic Ion Trap Fragmentation Pattern
 Cysteine residue mass: 160.03 Switch from monoisotopic to average mass at 5000
 Ions per window: 6.0 Extension Threshold: 0.15 Extension Number: 6
 Gaps: 1 Peak Width: 1.0 Data Threshold: 0.10 (30771) Ions per residue: 2.7
 Amino acids known to be present: *
 Amino acids known to be absent: *
 Unmodified N-terminus. Unmodified C-terminus.
 N-terminal Tag Mass:  0.00 C-terminal Tag Mass:  0.00 Sequence Tag: *0
 Edman data is not available.  AutoTag OFF
 Spectral Quality = 0.500000
Contiguous series of sequence ions defines a sequence of length 5

 Sequence                          Rank CombScr X-corr IntOnlyScr Quality
TG[220.1]LGHNHGR                      1  0.428   0.906   0.840     0.500

Search time: 0:00:05

          1st ranked  St Deviations  Average Wrong  Correct/Wrong
xcorr score     0.906      2.168        0.661         1.37
Int only score   0.840      1.764        0.669         1.26
Combined score   0.428      2.675        0.232         1.85
```

Figure B-6. Lutefisk report for the tryptic peptide TGPNLHGLFGR from the tryptic digest of cytochrome c. The *de novo*-derived (Lutefisk) sequence of TG[220.1]LGHNHGR provides additional validation of the database-derived (Sequest and Mascot) sequence (regions of sequence homology are underlined).

APPENDIX C

TROUBLESHOOTING THE *IN VIVO* MICRODIALYSIS-CLC-MS^2 SYSTEM

Table C-1 describes the symptoms, diagnosis and treatment for common problems in each dimension of the *in vivo* microdialysis-CLC-MS^2 system. This example is non-exhaustive and is only intended as a guide, in conjunction with selected references.

Table C-1. Symptoms, diagnosis and treatment for common problems in each dimension of the *in vivo* microdialysis-CLC-MS^2 system.

Dimension	Symptoms	Diagnosis	Treatment
In vivo microdialysis	Sweating of the probe.	Clog in outlet line, sample loop or waste line.	Replace all capillary lines.
	No measurements of expected peptides.	Incorrect placement of the probe.	Confirm coordinates and probe placement.
CLC	Poor flow.	Column clogged.	Etch the electrospray emitter in 50% HF for 10 s.
	Sample carryover.	Peptides at greater than nM levels are adsorbed to valve rotor.	Take apart the valve and clean the rotor and stator in IPA:10% HAc.
	Excessive gradient delay.	Poorly-made fittings or old valve rotor caused increased dead volume.	Remake the fittings and replace the valve rotor.
MS^2	Poor electrospray.	Potential and distance are not optimized.	Set the electrospray potential at 1.5 kV and the electrospray emitter less than 2 mm from the entrance of the QIT.
	Poor sensitivity.	Ion optics are not optimized.	Tune the ion optics automatically while infusing a solution of 18 nM NT 1-11 in 50% MeOH:1% HAc.
	Poor fragmentation.	Low He (g) pressure.	Replace He(g)-grade 5.

LIST OF REFERENCES

1. Krieger,D.T. Brain Peptides - What, Where, and Why. *Science* **222**, 975-985 (1983).

2. Hook,V. and Metz-Boutigue,M.H. Protein trafficking to chromaffin granules and proteolytic processing within regulated secretory vesicles of neuroendocrine chromaffin cells. *Chromaffin Cell: Trnsmitter Biosynthesis, Storage, Release, Actions, and Informatics* **971**, 397-405 (2002).

3. Kilpatrick,D.L., Jones,B.N., Lewis,R.V., Stern,A.S., Shively,J.E., and Udenfriend,S. An 18.2 Kdal Enkephalin-Containing Protein - An Intermediate in the Processing of Bovine Adrenal Proenkephalin. *Federation Proceedings* **41**, 1077 (1982).

4. Hughes,J., Smith,T.W., Kosterlitz,H.W., Fothergill,L.A., Morgan,B.A., and Morris,H.R. Identification of 2 Related Pentapeptides from Brain with Potent Opiate Agonist Activity. *Nature* **258**, 577-579 (1975).

5. Kennedy,R.T., Watson,C.J., Haskins,W.E., Powell,D.H., and Strecker,R.E. In vivo neurochemical monitoring by microdialysis and capillary separations. *Current Opinion in Chemical Biology* **6**, 659-665 (2002).

6. Haskins,W.E., Wang,Z.Q., Watson,C.J., Rostand,R.R., Witowski,S.R., Powell,D.H., and Kennedy,R.T. Capillary LC-MS2 at the attomole level for monitoring and discovering endogenous peptides in microdialysis samples collected in vivo. *Analytical Chemistry* **73**, 5005-5014 (2001).

7. Ungerstedt,U. and Hallstrom,A. Invivo Microdialysis - A New Approach to the Analysis of Neurotransmitters in the Brain. *Life Sciences* **41**, 861-864 (1987).

8. Tononi,G. and Cirelli,C. Modulation of brain gene expression during sleep and wakefulness: A review of recent findings. *Neuropsychopharmacology* **25**, S28-S35 (2001).

9. Kennedy,R.T. and Jorgenson,J.W. Preparation and Evaluation of Packed Capillary Liquid- Chromatography Columns with Inner Diameters from 20-Mu-M to 50-Mu-M. *Analytical Chemistry* **61**, 1128-1135 (1989).

10. Novotny,M.V. Microcolumn liquid chromatography in biochemical analysis. *High Resolution Separation and Analysis of Biological Macromolecules, Pt A* **270**, 101-133 (1996).

11. Andren,P.E., Emmett,M.R., and Caprioli,R.M. Micro-Electrospray - Zeptomole-Attomole Per Microliter Sensitivity for Peptides. *Journal of the American Society for Mass Spectrometry* **5**, 867-869 (1994).

12. Fenn,J.B., Mann,M., Meng,C.K., Wong,S.F., and Whitehouse,C.M. Electrospray Ionization-Principles and Practice. *Mass Spectrometry Reviews* **9**, 37-70 (1990).

13. Marina,A., Garcia,M.A., Albar,J.P., Yague,J., de Castro,J.A.L., and Vazquez,J. High-sensitivity analysis and sequencing of peptides and proteins by quadrupole ion trap mass spectrometry. *Journal of Mass Spectrometry* **34**, 17-27 (1999).

14. Andersen,J.S., Svensson,B., and Roepstorff,P. Electrospray ionization and matrix assisted laser desorption/ionization mass spectrometry: Powerful analytical tools in recombinant protein chemistry. *Nature Biotechnology* **14**, 449-457 (1996).

15. Yates,J.R. Mass spectrometry and the age of the proteome. *Journal of Mass Spectrometry* **33**, 1-19 (1998).

16. Yates,J.R. Database searching using mass spectrometry data. *Electrophoresis* **19**, 893-900 (1998).

17. Johnson,J.V. and Yost,R.A. Tandem Mass-Spectrometry for Trace Analysis. *Analytical Chemistry* **57**, A758-and (1985).

18. Maidment,N.T., Brumbaugh,D.R., Rudolph,V.D., Erdelyi,E., and Evans,C.J. Microdialysis of Extracellular Endogenous Opioid-Peptides from Rat-Brain Invivo. *Neuroscience* **33**, 549-557 (1989).

19. Lovelace,J.L., Kusmierz,J.J., and Desiderio,D.M. Analysis of Methionine Enkephalin in Human Pituitary by Multidimensional Reversed-Phase High-Performance Liquid- Chromatography, Radioreceptor Assay, Radioimmunoassay, Fast- Atom-Bombardment Mass-Spectrometry, and Mass-Spectrometry Mass-Spectrometry. *Journal of Chromatography-Biomedical Applications* **562**, 573-584 (1991).

20. Nilsson,C.L., Karlsson,G., Bergquist,J., Westman,A., and Ekman,R. Mass spectrometry of peptides in neuroscience. *Peptides* **19**, 781-789 (1998).

21. Shen,H., Lada,M.W., and Kennedy,R.T. Monitoring of met-enkephalin in vivo with 5-min temporal resolution using microdialysis sampling and capillary liquid chromatography with electrochemical detection. *Journal of Chromatography B* **704**, 43-52 (1997).

22. Shen,H., Witowski,S.R., Boyd,B.W., and Kennedy,R.T. Detection of peptides by precolumn derivatization with biuret reagent and preconcentration on capillary liquid chromatography columns with electrochemical detection. *Analytical Chemistry* **71**, 987-994 (1999).

23. Emmett,M.R., Andren,P.E., and Caprioli,R.M. Specific molecular mass detection of endogenously released neuropeptides using in vivo microdialysis mass spectrometry. *Journal of Neuroscience Methods* **62**, 141-147 (1995).

24. Andren,P.E. and Caprioli,R.M. Determination of extracellular release of neurotensin in discrete rat brain regions utilizing in vivo microdialysis/electrospray mass spectrometry. *Brain Research* **845**, 123-129 (1999).

25. Tiller,P.R., Cunniff,J.B., and Land,A.P. Advantages of full scan tandem mass spectrometry for quantitative analysis on a modern ion trap mass spectrometer: Increased selectivity and mixture analysis. *Rapid Communications in Mass Spectrometry* **11**, 1151-1153 (1997).

26. Li,L.Y.T., Campbell,D.A., Bennett,P.K., and Henion,J. Acceptance criteria for ultratrace HPLC-tandem mass spectrometry: Quantitative and qualitative determination of sulfonylurea herbicides in soil. *Analytical Chemistry* **68**, 3397-3404 (1996).

27. Hunt,D.F., Yates,J.R., Shabanowitz,J., Winston,S., and Hauer,C.R. Protein Sequencing by Tandem Mass-Spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 6233-6237 (1986).

28. Viklund,C., Ponten,E., Glad,B., Irgum,K., Horstedt,P., and Svec,F. "Molded" macroporous poly(glycidyl methacrylate-co- trimethylolpropane trimethacrylate) materials with fine controlled porous properties: Preparation of monoliths using photoinitiated polymerization. *Chemistry of Materials* **9**, 463-471 (1997).

29. Chen,J.R., Dulay,M.T., Zare,R.N., Svec,F., and Peters,E. Macroporous photopolymer frits for capillary electrochromatography. *Analytical Chemistry* **72**, 1224-1227 (2000).

30. Valaskovic,G.A., Kelleher,N.L., Little,D.P., Aaserud,D.J., and McLafferty,F.W. Attomole-Sensitivity Electrospray Source for Large-Molecule Mass-Spectrometry. *Analytical Chemistry* **67**, 3802-3805 (1995).

31. Snyder,L.R., Stadalius,M.A., and Quarry,M.A. Gradient Elution in Reversed-Phase Hplc Separation of Macromolecules. *Analytical Chemistry* **55**, 1412-and (1983).

32. Korner,R., Wilm,M., Morand,K., Schubert,M., and Mann,M. Nano electrospray combined with a quadrupole ion trap for the analysis of peptides and protein digests. *Journal of the American Society for Mass Spectrometry* **7**, 150-156 (1996).

33. Wieboldt,R., Campbell,D.A., and Henion,J. Quantitative liquid chromatographic tandem mass spectrometric determination of orlistat in plasma with a quadrupole ion trap. *Journal of Chromatography B* **708**, 121-129 (1998).

34. Hunt,D.F. Automated identification of peptides and proteins at the attomole level in complex mixtures by mass spectrometry. *Abstracts of Papers of the American Chemical Society* **222**, 1-TOXI (2001).

35. Hunt,D.F. Automated identification of peptides and proteins at the attomole level in complex mixtures by mass spectrometry. *Abstracts of Papers of the American Chemical Society* **221**, 151-ANYL (2001).

36. Hunt,D.F. Proteomics: Automated identification of peptides and proteins at the attomole level in complex mixtures by mass spectrometry. *Abstracts of Papers of the American Chemical Society* **219**, 249-ORGN (2000).

37. Settlage,R.E., Russo,P.S., Shabanowitz,J., and Hunt,D.F. A novel mu-ESI source for coupling capillary electrophoresis and mass spectrometry: Sequence determination of tumor peptides at the attomole level. *Journal of Microcolumn Separations* **10**, 281-285 (1998).

38. Yoshikawa,K., Williams,C., and Sabol,S.L. Rat-Brain Preproenkephalin Messenger-Rna - Cdna Cloning, Primary Structure, and Distribution in the Central Nervous- System. *Journal of Biological Chemistry* **259**, 4301-4308 (1984).

39. Fallon,J.H. and Leslie,F.M. Distribution of Dynorphin and Enkephalin Peptides in the Rat- Brain. *Journal of Comparative Neurology* **249**, 293-336 (1986).

40. Merchenthaler,I., Maderdrut,J.L., Altschuler,R.A., and Petrusz,P. Immunocytochemical Localization of Proenkephalin-Derived Peptides in the Central-Nervous-System of the Rat. *Neuroscience* **17**, 325-and (1986).

41. Hoover,B.R. and Marshall,J.F. Further characterization of preproenkephalin mRNA-containing cells in the rodent globus pallidus. *Neuroscience* **111**, 111-125 (2002).

42. Hoover,B.R. and Marshall,J.F. Population characteristics of preproenkephalin mRNA-containing neurons in the globus pallidus of the rat. *Neuroscience Letters* **265**, 199-202 (1999).

43. Goumon,Y., Lugardon,K., Gadroy,P., Strub,J.M., Welters,I.D., Stefano,G.B., Aunis,D., and Metz-Boutigue,M.H. Processing of proenkephalin-A in bovine chromaffin cells - Identification of natural derived fragments by N-terminal sequencing and matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of Biological Chemistry* **275**, 38355-38362 (2000).

44. Baird,A., Klepper,R., and Ling,N. Invitro and Invivo Evidence That the C-Terminus of Preproenkephalin-A Circulates As An 8500-Dalton Molecule. *Proceedings of the Society for Experimental Biology and Medicine* **175**, 304-308 (1984).

45. Liston,D., Patey,G., Rossier,J., Verbanck,P., and Vanderhaeghen,J.J. Processing of Proenkephalin Is Tissue-Specific. *Science* **225**, 734-737 (1984).

46. Kilpatrick,D.L., Taniguchi,T., Jones,B.N., Stern,A.S., Shively,J.E., Hullihan,J., Kimura,S., Stein,S., and Udenfriend,S. A Highly Potent 3200-Dalton Adrenal Opioid Peptide That Contains Both A [Met]-Enkephalin and [Leu]Enkephalin Sequence. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **78**, 3265-3268 (1981).

47. Zhang,H., Stoeckli,M., Andren,P.E., and Caprioli,R.M. Combining solid-phase preconcentration, capillary electrophoresis and off-line matrix-assisted laser desorption/ionization mass spectrometry: Intracerebral metabolic processing of peptide E in vivo. *Journal of Mass Spectrometry* **34**, 377-383 (1999).

48. Fricker,L.D., Berman,Y.L., Leiter,E.H., and Devi,L.A. Carboxypeptidase E activity is deficient in mice with the fat mutation - Effect on peptide processing. *Journal of Biological Chemistry* **271**, 30619-30624 (1996).

49. Bures,E.J., Courchesne,P.L., Douglass,J., Chen,K., Davis,M.T., Jones,M.T., Jones,M.D., McGinley,M.D., Robinson,J.H., Spahr,C.S., Sun,J.L., Wahl,R.C., and Patterson,S.D. Identification of incompletely processed potential Carboxypeptidase E substrates from CpEfat/CpEfat mice. *Proteomics* **1**, 79-92 (2001).

50. Che,F.Y., Yan,L., Li,H., Mzhavia,N., Devi,L.A., and Fricker,L.D. Identification of peptides from brain and pituitary of Cpe(fat)/Cpe(fat) mice. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 9971-9976 (2001).

51. Skold,K., Svensson,M., Kaplan,A., Bjorkesten,L., Astrom,J., and Andren,P.E. A neuroproteomic approach to targeting neuropeptides in the brain. *Proteomics* **2**, 447-454 (2002).

52. Mills,J. and Reiner,P.B. Regulation of amyloid precursor protein cleavage. *Journal of Neurochemistry* **72**, 443-460 (1999).

53. Caprioli,R.M., Moore,W.T., Martin,M., Dague,B.B., Wilson,K., and Moring,S. Coupling Capillary Zone Electrophoresis and Continuous-Flow Fast Atom Bombardment Mass-Spectrometry for the Analysis of Peptide Mixtures. *Journal of Chromatography* **480**, 247-257 (1989).

54. Andren,P.E., Emmett,M.R., and Caprioli,R.M. Specific molecular mass detection of endogenously released neuropeptides using in vivo microdialysis mass spectrometry. *Journal of Neurochemistry* **66**, S66 (1996).

55. Yates,J.R., Morgan,S.F., Gatlin,C.L., Griffin,P.R., and Eng,J.K. Method to compare collision-induced dissociation spectra of peptides: Potential for library searching and subtractive analysis. *Analytical Chemistry* **70**, 3557-3565 (1998).

56. Ducret,A., Van Oostveen,I., Eng,J.K., Yates,J.R., and Aebersold,R. High throughput protein characterization by automated reverse- phase chromatography electrospray tandem mass spectrometry. *Protein Science* **7**, 706-719 (1998).

57. Peng,J.M., Elias,J.E., Thoreen,C.C., Licklider,L.J., and Gygi,S.P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: The yeast proteome. *Journal of Proteome Research* **2**, 43-50 (2003).

58. MacCoss,M.J., Wu,C.C., and Yates,J.R. Probability-based validation of protein identifications using a modified SEQUEST algorithm. *Analytical Chemistry* **74**, 5593-5599 (2002).

59. Perkins,D.N., Pappin,D.J.C., Creasy,D.M., and Cottrell,J.S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **20**, 3551-3567 (1999).

60. Taylor,J.A. and Johnson,R.S. Implementation and uses of automated de novo peptide sequencing by tandem mass spectrometry. *Analytical Chemistry* **73**, 2594-2604 (2001).

61. Horwitz,W., Kamps,L.R., and Boyer,K.W. Quality Assurance in the Analysis of Foods for Trace Constituents. *Journal of the Association of Official Analytical Chemists* **63**, 1344-1354 (1980).

62. Simpson,R.J., Connolly,L.M., Eddes,J.S., Pereira,J.J., Moritz,R.L., and Reid,G.E. Proteomic analysis of the human colon carcinoma cell line (LIM 1215): Development of a membrane protein database. *Electrophoresis* **21**, 1707-1732 (2000).

63. McLafferty,F.W., Stauffer,D.A., Loh,S.Y., and Wesdemiotis,C. Unknown identification using reference mass spectra. Quality evaluation of databases. *Journal of the American Society for Mass Spectrometry* **10**, 1229-1240 (1999).

64. Hough,J.M., Haney,C.A., Voyksner,R.D., and Bereman,R.D. Evaluation of electrospray transport CID for the generation of searchable libraries. *Analytical Chemistry* **72**, 2265-2270 (2000).

65. Haskins,W.E., Wang,Z.Q., Watson,C.J., Rostand,R.R., Witowski,S.R., Powell,D.H., and Kennedy,R.T. Capillary LC-MS2 at the attomole level for monitoring and discovering endogenous peptides in microdialysis samples collected in vivo. *Analytical Chemistry* **73**, 5005-5014 (2001).

66. Choudhary,J.S., Blackstock,W.P., Creasy,D.M., and Cottrell,J.S. Interrogating the human genome using uninterpreted mass spectrometry data. *Proteomics* **1**, 651-667 (2001).

67. Myers,R.D. Neuroactive Peptides - Unique Phases in Research on Mammalian Brain Over 3 Decades. *Peptides* **15**, 367-381 (1994).

68. Rawlings,N.D., O'Brien,E., and Barrett,A.J. MEROPS: the protease database. *Nucl. Acids. Res.* **30**, 343-346 (2002).

69. Lopez-Otin,C. and Overall,C.M. Protease degradomics: A new challenge for proteomics. *Nature Reviews Molecular Cell Biology* **3**, 509-519 (2002).

70. Brooks,D.A. Protein processing: A role in the pathophysiology of genetic disease. *Febs Letters* **409**, 115-120 (1997).

71. Seidah,N.G. and Chretien,M. Eukaryotic protein processing: endoproteolysis of precursor proteins. *Current Opinion in Biotechnology* **8**, 602-607 (1997).

72. Rouille,Y., Duguay,S.J., Lund,K., Furuta,M., Gong,Q.M., Lipkind,G., Oliva,A.A., Chan,S.J., and Steiner,D.F. Proteolytic Processing Mechanisms in the Biosynthesis of Neuroendocrine Peptides - the Subtilisin-Like Proprotein Convertases. *Frontiers in Neuroendocrinology* **16**, 322-361 (1995).

73. Devi,L. Consensus Sequence for Processing of Peptide Precursors at Monobasic Sites. *Febs Letters* **280**, 189-194 (1991).

74. Cupo,A., Eybalin,M., Patey,G., Rossier,J., and Jarry,T. Quantitation and Localization of Met-Enkephalin-Arg-Gly-Leu in Rat-Brain Using Highly Sensitive Antibodies. *Neuropeptides* **4**, 389-401 (1984).

75. Patey,G., Cupo,A., Mazarguil,H., Morgat,J.L., and Rossier,J. Release of Proenkephalin-Derived Opioid-Peptides from Rat Striatum Invitro and Their Rapid Degradation. *Neuroscience* **15**, 1035-1044 (1985).

76. Patey,G., Cupo,A., Chaminade,M., Morgat,J.L., and Rossier,J. Release of the Heptapeptide Met-Enkephalin-Arg6-Phe7 and of the Octapeptide Met-Enkephalin-Arg6-Gly7-Leu8 from Rat Striatum Invitro and Their Rapid Inactivation. *Life Sciences* **33**, 117-120 (1983).

77. Patey,G., Cupo,A., Giraud,P., Chaminade,M., and Rossier,J. The Heptapeptide Met-Enkephalin-Arg6, Phe7 Is Released from Rat Striatum Invitro by High Potassium. *Neuroscience Letters* **37**, 267-271 (1983).

78. Fayada,C., Guibert,B., Patey,G., Cupo,A., Chaminade,M., Rossier,J., and Leviel,V. Release of [Met]Enkephalin in the Central Nucleus of the Amygdala Is Increased by Application of Potassium in the Substantia-Nigra. *Neuropeptides* **9**, 9-17 (1987).

79. Wang,H.P. and Dass,C. Characterization of bioactive peptides in bovine adrenal medulla by a combination of fast HPLC and ESI-MS. *Peptides* **23**, 2143-2150 (2002).

80. Hook,V. and Metz-Boutigue,M.H. Protein trafficking to chromaffin granules and proteolytic processing within regulated secretory vesicles of neuroendocrine chromaffin cells. *Chromaffin Cell: Trnsmitter Biosynthesis, Storage, Release, Actions, and Informatics* **971**, 397-405 (2002).

81. Lembo,P.M.C., Grazzini,E., Groblewski,T., O'Donnell,D., Roy,M.O., Zhang,J., Hoffert,C., Cao,J., Schmidt,R., Pelletier,M., Labarre,M., Gosselin,M., Fortin,Y., Banville,D., Shen,S.H., Strom,P., Payza,K., Dray,A., Walker,P., and Ahmad,S. Proenkephalin A gene products activate a new family of sensory neuron-specific GPCRs. *Nature Neuroscience* **5**, 201-209 (2002).

82. Goumon,Y., Lugardon,K., Gadroy,P., Strub,J.M., Welters,I.D., Stefano,G.B., Aunis,D., and Metz-Boutigue,M.H. Processing of proenkephalin-A in bovine chromaffin cells - Identification of natural derived fragments by N-terminal sequencing and matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of Biological Chemistry* **275**, 38355-38362 (2000).

83. Condamine,E., Leprince,J., Suaudeau,C., Mayer,C., Davoust,D., Costentin,J., and Vaudry,H. The proenkephalin A-processing product peptide E, which encompasses two enkephalin sequences, has a much lower opioid activity than beta-endorphin. *Peptides* **20**, 865-871 (1999).

84. Hook,V.Y.H., Noctor,S., Sei,C.A., Toneff,T., Yasothornsrikul,S., and Kang,Y.H. Evidence for functional localization of the proenkephalin- processing enzyme, prohormone thiol protease, to secretory vesicles of chromaffin cells. *Endocrinology* **140**, 3744-3754 (1999).

85. Yasothornsrikul,S., Aaron,W., Toneff,T., and Hook,V.Y.H. Evidence for the proenkephalin processing enzyme prohormone thiol protease (PTP) as a multicatalytic cysteine protease complex: Activation by glutathione localized to secretory vesicles. *Biochemistry* **38**, 7421-7430 (1999).

86. Goumon,Y., Lugardon,K., Kieffer,B., Lefevre,J.F., Van Dorsselaer,A., Aunis,D., and Metz-Boutigue,M.H. Characterization of antibacterial COOH-terminal proenkephalin- A-derived peptides (PEAP) in infectious fluids - Importance of enkelytin, the antibacterial PEAP(209-237) secreted by stimulated chromaffin cells. *Journal of Biological Chemistry* **273**, 29847-29856 (1998).

87. Krieger,T.J., Mendemueller,L., and Hook,V.Y.H. Prohormone Thiol Protease and Enkephalin Precursor Processing - Cleavage at Dibasic and Monobasic Sites. *Journal of Neurochemistry* **59**, 26-31 (1992).

88. Wilson,S.P., Corcoran,J.J., and Kirshner,N. Comparative Incorporation of Proenkephalin-Derived Peptides, Chromogranin-A, and Dopamine Beta-Hydroxylase Into Chromaffin Vesicles. *Journal of Neurochemistry* **57**, 870-875 (1991).

89. Mindroiu,T., Carretero,O.A., Walz,D., and Scicli,A.G. Tissue Kallikrein Processes Small Proenkephalin Peptides. *Biochimica et Biophysica Acta* **1076**, 9-14 (1991).

90. Ase,A., Vindrola,O., Aloyz,R., Saravia,F., Finkielman,S., and Nahmod,V. Association of Proenkephalin-Derived High-Molecular-Weight Precursors with Microsomes and Granules Secreted by the Bovine Adrenal-Medulla. *Medicina-Buenos Aires* **47**, 606-607 (1987).

91. Boarder,M.R., Evans,C., Adams,M., Erdelyi,E., and Barchas,J.D. Peptide-e and Its Products, Bam-18 and Leu-Enkephalin, in Bovine Adrenal-Medulla and Cultured Chromaffin Cells - Release in Response to Stimulation. *Journal of Neurochemistry* **49**, 1824-1832 (1987).

92. Watkinson,A. and Dockray,G.J. Isolation and Characterization of N-Terminally Extended Met- Enkephalin Arg6Phe7 Peptides from Bovine Adrenal-Medulla and Porcine Digestive-Tract. *Regulatory Peptides* **18**, 350 (1987).

93. Dumont,M. and Lemaire,S. Molecular-Forms and Processing of Dynorphin-Related Peptides in Bovine Adrenal-Medulla - Comparison with Enkephalins and Precursors. *Federation Proceedings* **46**, 1448 (1987).

94. Birch,N.P. and Christie,D.L. Characterization of the Molecular-Forms of Proenkephalin in Bovine Adrenal-Medulla and Rat Adrenal, Brain, and Spinal-Cord with A Site-Directed Antiserum. *Journal of Biological Chemistry* **261**, 2213-2221 (1986).

95. Emmett,M.R., Andren,P.E., and Caprioli,R.M. Specific molecular mass detection of endogenously released neuropeptides using in vivo microdialysis mass spectrometry. *Journal of Neuroscience Methods* **62**, 141-147 (1995).

96. Bergasa,N.V., Mohajer,B., Maisonneuve,I.M., Ho,A., Maidment,N.T., Olive,F., Gunduz,M., and Kreek,M.J. Basal total opioid peptide release in the stratum of rats with cholestasis from bile duct resection: A study by the use of in vivo microdialysis. *Life Sciences* **61**, 1169-1175 (1997).

97. Haskins,W.E., Wang,Z.Q., Watson,C.J., Rostand,R.R., Witowski,S.R., Powell,D.H., and Kennedy,R.T. Capillary LC-MS2 at the attomole level for monitoring and discovering endogenous peptides in microdialysis samples collected in vivo. *Analytical Chemistry* **73**, 5005-5014 (2001).

98. Lapeyre,S., Mauborgne,A., Becker,C., Benoliel,J.J., Cesselin,F., Hamon,M., and Bourgoin,S. Subcutaneous formalin enhances outflow of met-enkephalin- and cholecystokinin-like materials in the rat nucleus accumbens. *Naunyn-Schmiedebergs Archives of Pharmacology* **363**, 399-406 (2001).

99. Yamamoto,Y., Sokawa,Y., and Maekawa,S. Biochemical evidence for the presence of NAP-22, a novel acidic calmodulin binding protein, in the synaptic vesicles of rat brain. *Neuroscience Letters* **224**, 127-130 (1997).

100. Rawlings,N.D., O'Brien,E., and Barrett,A.J. MEROPS: the protease database. *Nucl. Acids. Res.* **30**, 343-346 (2002).

101. Vernigora,A.N. and Gengin,M.T. Proteolytic enzymes: Subcellular location, properties, and involvement in neuropeptide turnover. A review. *Biochemistry-Moscow* **61**, 555-564 (1996).

102. Turgeon,V.L. and Houenou,L.J. The role of thrombin-like (serine) proteases in the development, plasticity and pathology of the nervous system. *Brain Research Reviews* **25**, 85-95 (1997).

103. Breslin,M.B., Lindberg,I., Benjannet,S., Mathis,J.P., Lazure,C., and Seidah,N.G. Differential Processing of Proenkephalin by Prohormone Convertase-1(3) and Convertase-2 and Furin. *Journal of Biological Chemistry* **268**, 27084-27093 (1993).

104. Lindberg,I., Ahn,S.C., and Breslin,M.B. Cellular-Distributions of the Prohormone Processing Enzymes Pc1 and Pc2. *Molecular and Cellular Neuroscience* **5**, 614-622 (1994).

105. Fayada,C., Guibert,B., Patey,G., Cupo,A., Chaminade,M., Rossier,J., and Leviel,V. Release of [Met]Enkephalin in the Central Nucleus of the Amygdala Is Increased by Application of Potassium in the Substantia-Nigra. *Neuropeptides* **9**, 9-17 (1987).

106. Patey,G., Cupo,A., Chaminade,M., Morgat,J.L., and Rossier,J. Release of the Heptapeptide Met-Enkephalin-Arg6-Phe7 and of the Octapeptide Met-Enkephalin-Arg6-Gly7-Leu8 from Rat Striatum Invitro and Their Rapid Inactivation. *Life Sciences* **33**, 117-120 (1983).

107. Patey,G., Liston,D., and Rossier,J. Characterization of New Enkephalin-Containing Peptides in the Adrenal-Medulla by Immunoblotting. *Febs Letters* **172**, 303-308 (1984).

108. Patey,G., Cupo,A., Mazarguil,H., Morgat,J.L., and Rossier,J. Release of Proenkephalin-Derived Opioid-Peptides from Rat Striatum Invitro and Their Rapid Degradation. *Neuroscience* **15**, 1035-1044 (1985).

109. Cupo,A., Eybalin,M., Patey,G., Rossier,J., and Jarry,T. Quantitation and Localization of Met-Enkephalin-Arg-Gly-Leu in Rat-Brain Using Highly Sensitive Antibodies. *Neuropeptides* **4**, 389-401 (1984).

110. Bodnar,R.J. and Hadjimarkou,M.M. Endogenous opiates and behavior: 2001. *Peptides* **23**, 2307-2365 (2002).

111. Jones,B.N., Shively,J.E., Kilpatrick,D.L., Kojima,K., and Udenfriend,S. Enkephalin Biosynthetic-Pathway - A 5300-Dalton Adrenal Polypeptide That Terminates at Its Cooh End with the Sequence [Met]Enkephalin-Arg-Gly-Leu-Cooh. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **79**, 1313-1315 (1982).

112. Bowser,M.T. and Kennedy,R.T. In vivo monitoring of amine neurotransmitters using microdialysis with on-line capillary electrophoresis. *Electrophoresis* **22**, 3668-3676 (2001).

113. Watanabe,M., Maemura,K., Kanbara,K., Tamayama,T., and Hayasaki,H. GABA and GABA receptors in the central nervous system and other organs. *International Review of Cytology - A Survey of Cell Biology, Vol 213* **213**, 1-47 (2002).

114. Stein,V. and Nicoll,R.A. GABA generates excitement. *Neuron* **37**, 375-378 (2003).

115. Saransaari,P. and Oja,S.S. Release of endogenous glutamate, aspartate, GABA, and taurine from hippocampal slices from adult and developing mice under cell-damaging conditions. *Neurochemical Research* **23**, 563-570 (1998).

116. Petroff,O.A.C. GABA and glutamate in the human brain. *Neuroscientist* **8**, 562-573 (2002).

117. Owens,D.F. and Kriegstein,A.R. Is there more to GABA than synaptic inhibition? *Nature Reviews Neuroscience* **3**, 715-727 (2002).

118. Osborne,P.G., OConnor,W.T., Drew,K.L., and Ungerstedt,U. An Invivo Microdialysis Characterization of Extracellular Dopamine and Gaba in Dorsolateral Striatum of Awake Freely Moving and Halothane Anesthetized Rats. *Journal of Neuroscience Methods* **34**, 99-105 (1990).

119. Nong,Y., Sorenson,E.M., and Chiappinelli,V.A. Opioid receptor activation attenuates nicotinic enhancement of spontaneous GABA release in lateral spiriform nucleus of the chick. *Brain Research* **961**, 45-52 (2003).

120. Grobin,A.C. and Deutch,A.Y. Dopaminergic regulation of extracellular gamma-aminobutyric acid levels in the prefrontal cortex of the rat. *Journal of Pharmacology and Experimental Therapeutics* **285**, 350-357 (1998).

121. Slemmon,J.R., Feng,B.B., and Erhardt,J.A. Small proteins that modulate calmodulin-dependent signal transduction - Effects of PEP-19, neuromodulin, and neurogranin on enzyme activation and cellular homeostasis. *Molecular Neurobiology* **22**, 99-113 (2000).

122. Mons,N., Enderlin,V., Jaffard,R., and Higueret,P. Selective age-related changes in the PKC-sensitive, calmodulin- binding protein, neurogranin, in the mouse brain. *Journal of Neurochemistry* **79**, 859-867 (2001).

123. Li,G.L., Farooque,M., Lewen,A., Lennmyr,F., Holtz,A., and Olsson,Y. MAP2 and neurogranin as markers for dendritic lesions in CNS injury. An immunohistochemical study in the rat. *Apmis* **108**, 98-106 (2000).

124. Gerendasy,D.D. and Sutcliffe,J.G. RC3/neurogranin, a postsynaptic calpacitin for setting the response threshold to calcium influxes. *Molecular Neurobiology* **15**, 131-163 (1997).

125. Chakravarthy,B., Morley,P., and Whitfield,J. Ca2+-calmodulin and protein kinase Cs: a hypothetical synthesis of their conflicting convergences on shared substrate domains. *Trends in Neurosciences* **22**, 12-16 (1999).

126. Baudier,J., Deloulme,J.C., Vandorsselaer,A., Black,D., and Matthes,H.W.D. Purification and Characterization of A Brain-Specific Protein- Kinase-C Substrate, Neurogranin (P17) - Identification of A Consensus Amino-Acid-Sequence Between Neurogranin and Neuromodulin (Gap43) That Corresponds to the Protein-Kinase-C Phosphorylation Site and the Calmodulin-Binding Domain. *Journal of Biological Chemistry* **266**, 229-237 (1991).

127. Maekawa,S., Maekawa,M., Hattori,S., and Nakamura,S. Purification and Molecular-Cloning of A Novel Acidic Calmodulin-Binding Protein from Rat-Brain. *Journal of Biological Chemistry* **268**, 13703-13709 (1993).

128. Iino,S., Kobayashi,S., and Maekawa,S. Immunohistochemical localization of a novel acidic calmodulin- binding protein, NAP-22, in the rat brain. *Neuroscience* **91**, 1435-1444 (1999).

129. Gegelashvili,G. and Schousboe,A. High affinity glutamate transporters: Regulation of expression and activity. *Molecular Pharmacology* **52**, 6-15 (1997).

130. Mitrovic,A.D., Amara,S.G., Johnston,G.A.R., and Vandenberg,R.J. Identification of functional domains of the human glutamate transporters EAAT1 and EAAT2. *Journal of Biological Chemistry* **273**, 14698-14706 (1998).

131. Raiteri,L., Stigliani,S., Zedda,L., Raiteri,M., and Bonanno,G. Multiple mechanisms of transmitter release evoked by 'pathologically' elevated extracellular [K+]: involvement of transporter reversal and mitochondrial calcium. *Journal of Neurochemistry* **80**, 706-714 (2002).

132. Phillis,J.W., Ren,J., and O'Regan,M.H. Transporter reversal as a mechanism of glutamate release from the ischemic rat cerebral cortex: studies with DL-threo-beta- benzyloxyaspartate. *Brain Research* **868**, 105-112 (2000).

133. RodriguezSanchez,P., TejeroDiez,P., and DiezGuerra,F.J. Glutamate stimulates neurogranin phosphorylation in cultured rat hippocampal neurons. *Neuroscience Letters* **221**, 137-140 (1997).

134. Chen,S.J., Sweatt,J.D., and Klann,E. Enhanced phosphorylation of the postsynaptic protein kinase C substrate RC3/neurogranin during long-term potentiation. *Brain Research* **749**, 181-187 (1997).

135. Neunerjehle,M., Rhyner,T.A., and Borbely,A.A. Sleep-Deprivation Differentially Alters the Messenger-Rna and Protein-Levels of Neurogranin in Rat-Brain. *Brain Research* **685**, 143-153 (1995).

136. Wang,K.K.W. Calpain and caspase: can you tell the difference? *Trends in Neurosciences* **23**, 20-26 (2000).

137. Blomback,B., Blomback,M., and Grondahl,N.J. Studies on Fibrinopeptides from Mammals. *Acta Chemica Scandinavica* **19**, 1789-and (1965).

138. Blomback,B., Blomback,M., Grondahl,N.J., Guthrie,C., and Hinton,M. Studies on Fibrinopeptides from Primates. *Acta Chemica Scandinavica* **19**, 1788-and (1965).

139. Tegernil,A.C. and Blomback,B. Degradataion of Human Fibrinopeptides in Serum. *Acta Chemica Scandinavica* **21**, 307-and (1967).

140. Ceriello,A. Coagulation Activation in Diabetes-Mellitus - the Role of Hyperglycemia and Therapeutic Prospects. *Diabetologia* **36**, 1119-1125 (1993).

141. Rickles,F.R., Levine,M., and Edwards,R.L. Hemostatic Alterations in Cancer-Patients. *Cancer and Metastasis Reviews* **11**, 237-248 (1992).

142. Mackie,E.J., Pagel,C.N., Smith,R., de Niese,M.R., Song,S.J., and Pike,R.N. Protease-activated receptors: A means of converting extracellular proteolysis into intracellular signals. *Iubmb Life* **53**, 277-281 (2002).

143. Wang,H. and Reiser,G. Thrombin signaling in the brain: The role of protease-activated receptors. *Biological Chemistry* **384**, 193-202 (2003).

144. McLafferty,F.W., Zhang,M.Y., Stauffer,D.B., and Loh,S.Y. Comparison of algorithms and databases for matching unknown mass spectra. *Journal of the American Society for Mass Spectrometry* **9**, 92-95 (1998).

145. McLafferty,F.W., Hirota,A., and Barbalas,M.P. Library of Collisional Activation Mass-Spectra. *Organic Mass Spectrometry* **15**, 327-328 (1980).

146. Strecker,R.E., Sharp,T., Brundin,P., Zetterstrom,T., Ungerstedt,U., and Bjorklund,A. Autoregulation of Dopamine Release and Metabolism by Intrastriatal Nigral Grafts As Revealed by Intracerebral Dialysis. *Neuroscience* **22**, 169-178 (1987).

147. Davidsson,P., Paulson,L., Hesse,C., Blennow,K., and Nilsson,C.L. Proteome studies of human cerebrospinal fluid and brain tissue using a preparative two-dimensional electophoresis approach prior to mass spectrometry. *Proteomics* **1**, 444-452 (2001).

148. Skold,K., Svensson,M., Kaplan,A., Bjorkesten,L., Astrom,J., and Andren,P.E. A neuroproteomic approach to targeting neuropeptides in the brain. *Proteomics* **2**, 447-454 (2002).

149. Che,F.Y., Yan,L., Li,H., Mzhavia,N., Devi,L.A., and Fricker,L.D. Identification of peptides from brain and pituitary of Cpe(fat)/Cpe(fat) mice. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 9971-9976 (2001).

150. McLafferty,F.W. Tandem mass spectrometric analysis of complex biological mixtures. *International Journal of Mass Spectrometry* **212**, 81-87 (2001).

151. Loh,S.Y. and McLafferty,F.W. Exact-Mass Probability Based Matching of High-Resolution Unknown Mass-Spectra. *Analytical Chemistry* **63**, 546-550 (1991).

152. Stauffer,D.B. and McLafferty,F.W. Molecular-Weight Adjustment to Probability Based Matching of Unknown Mass-Spectra. *Organic Mass Spectrometry* **21**, 313-315 (1986).

153. Atwater,B.L., Stauffer,D.B., McLafferty,F.W., and Peterson,D.W. Reliability Ranking and Scaling Improvements to the Probability Based Matching System for Unknown Mass-Spectra. *Analytical Chemistry* **57**, 899-903 (1985).

154. Stauffer,D.B., McLafferty,F.W., Ellis,R.D., and Peterson,D.W. Probability-Based-Matching Algorithm with Forward Searching Capabilities for Matching Unknown Mass-Spectra of Mixtures. *Analytical Chemistry* **57**, 1056-1060 (1985).

155. Pesyna,G.M., Venkataraghavan,R., Dayringer,H.E., and McLafferty,F.W. Probability Based Matching System Using A Large Collection of Reference Mass-Spectra. *Analytical Chemistry* **48**, 1362-1368 (1976).

156. Mclaffer,F.W., Hertel,R.H., and Villwock,R.D. Computer Identification of Mass-Spectra .6. Probability Based Matching of Mass-Spectra - Rapid Identification of Specific Compounds in Mixtures. *Organic Mass Spectrometry* **9**, 690-702 (1974).

157. Stahle,L., Segersvard,S., and Ungerstedt,U. A Comparison Between 3 Methods for Estimation of Extracellular Concentrations of Exogenous and Endogenous Compounds by Microdialysis. *Journal of Pharmacological Methods* **25**, 41-52 (1991).

BIOGRAPHICAL SKETCH

William E. Haskins was born April 13, 1973, in Camp Hill, Pennsylvania. He obtained a Bachelor of Science degree in chemistry from New Mexico State University in December 1995. He worked as an analyst in the Electronic Chemicals Division of Ashland Chemical in Dallas, Texas from 1995 to 1996. Then, he began graduate research with Professor Shuming Nie in the Chemistry Department at Indiana University where he obtained a Master of Science degree in bioanalytical chemistry. In August 1998, he continued his graduate studies with Dr. Robert T. Kennedy in the Chemistry Department at the University of Florida. He completed his research in August 2003, obtaining a Doctor of Philosophy degree in bioanalytical chemistry.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert T. Kennedy, Chair
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

David H. Powell
Scientist of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Richard A. Yost
Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Benjamin Horenstein
Associate Professor of Chemistry

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Satya P. Kalra
Professor of Neuroscience

This dissertation was submitted to the Graduate Faculty of the Department of Chemistry in the College of Liberal Arts and Sciences and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 2003

Dean, Graduate School